

UNIVERSIDAD AUTÓNOMA DE MADRID

ESCUELA POLITÉCNICA SUPERIOR



**MÁSTER UNIVERSITARIO EN INVESTIGACIÓN E
INNOVACIÓN EN TIC (i²-TIC)**

TRABAJO FIN DE MASTER

**ANÁLISIS Y ESTUDIO DE RONQUIDOS PARA LA
DETECCIÓN DE TRASTORNOS RESPIRATORIOS
DURANTE EL SUEÑO**

FRANCISCO RODRÍGUEZ GONZÁLEZ

SEPTIEMBRE 2015

RESUMEN

El Síndrome de Apnea-Hipoapnea del Sueño (SAHS) es un trastorno bastante común, con una prevalencia del 2% al 7% de la población. Se estima que el 90% de los enfermos no están diagnosticados, ya que la prueba diagnóstica de Polisomnografía es cara y limitada por el número de camas disponibles en la Unidad del Sueño de los Hospitales. El SAHS presenta un alto impacto social y económico debido a la falta de un descanso adecuado durante el sueño que produce estados de somnolencia durante la vigilia que pueden provocar descenso en la productividad en el trabajo y accidentes en el manejo de maquinaria y vehículos. Así mismo, la situación de hipoxia durante el sueño puede inducir diversas enfermedades circulatorias, neurocognitivas y metabólicas en función de su severidad. Por lo tanto, el poder disponer de un sistema de recomendación de diagnóstico de SAHS de bajo coste, sencillo y que alcance a un gran porcentaje de la población sería de gran utilidad para los Servicios Médicos implicados en el diagnóstico y tratamiento del SAHS.

En este Trabajo de Fin de Master se ha propuesto un método de análisis de los sonidos respiratorios durante el sueño (silencio – respiración – ronquido), registrados fuera de un entorno controlado de laboratorio. Los registros nocturnos fueron realizados por voluntarios en sus domicilios, utilizando sus propios *Smartphones* o un grabador MP3, sin supervisión directa. De los registros nocturnos, se extrajeron segmentos de audio con los sonidos de actividad respiratoria que contuvieran poco ruido de fondo (ruidos de vehículos, conversaciones, cantos de pájaros, sonidos de mecanismos), siempre que fuera posible. Estos segmentos se dividieron a su vez en ventanas de 100 ms, que fueron clasificadas manualmente en tres clases: Silencio – Respiración – Ronquido. Así mismo, para cada ventana se extrajeron cuatro características a partir de la señal de audio: Zero Crossing Rate, Short Time Energy, Coeficiente de autocorrelación normalizado a 1 ms de retardo y Primer Coeficiente predictor de análisis LPC (Linear Prediction Coding) de 12 polos.

Se ha diseñado un método para calcular una recomendación de diagnóstico de SAHS basado en las cuatro características de audio expuestas. Se han evaluado distintos métodos de clasificación, buscando el que mejor rendimiento ofreciera en este problema de clasificación. En este caso, un modelo de RandomForest aplicado a las características extraídas de la señal de audio de los registros nocturnos proporciona una clasificación correcta del diagnóstico SAHS / No-SAHS del 97.33%, con una sensibilidad del 0.973, una especificidad de 0.973 y una AUC (área bajo la curva ROC) de 0.996.

Así mismo, las características de audio registradas permiten una clasificación *on-line* de los eventos respiratorios, ya sean Respiración/Silencio/Ronquido, o No-ronquido/Ronquido. En ambos casos, el modelo de clasificación con mejor rendimiento de los evaluados es RandomForest. La clasificación correcta de Respiración/Silencio/ronquido se produce con un 85.26% de éxito, con una sensibilidad de 0.853, una especificidad de 0.892 y una AUC de 0.944. La clasificación correcta en Ronquido/No-ronquido se produce con un 92.84% de éxito, una sensibilidad de 0.928, una especificidad de 0.859 y AUC de 0.959.

Adicionalmente, se ha comprobado que la incorporación de datos biométricos de los voluntarios recogidos mediante formularios STOP-BANG al modelo de clasificación mejora los resultados. En la clasificación de eventos respiratorios Respiración/Silencio/Ronquido se obtiene un 87.09% de éxito con una sensibilidad de 0.871, una especificidad de 0.903 y AUC de 0.953 al añadir el peso, altura y edad del voluntario. En la clasificación de eventos No-ronquido/Ronquido, se mejora la clasificación con RandomForest hasta un 93.54% de éxito,

con especificidad de 0.935, sensibilidad de 0.867 y AUC de 0.962 al añadir también peso, altura y edad. Respecto a la clasificación de diagnóstico SAHS/No-SAHS, usando RandomForest de obtiene un 97.95% de éxito, con sensibilidad de 0.98, especificidad de 0.979 y AUC de 0.997 al añadir el género del paciente. Debido al bajo número de voluntarios (7), no se pudieron evaluar todos los datos biométricos recogidos, por lo que habrá que seguir estudiando su aporte en futuros trabajos cuando se disponga de una población de experimentación mayor.

Los resultados de diagnóstico de SAHS constituyen un primer paso para el desarrollo futuro de una aplicación de Smartphone que puede ser distribuida como una herramienta de cribado entre la población, de manera que analice los sonidos respiratorios durante el sueño y realice una recomendación de visita al médico especialista si se calcula una alta probabilidad de sufrir SAHS.

Los resultados de la clasificación de eventos respiratorios abren la puerta a desarrollar sistemas que estimulen al usuario cuando ronca para intentar evitar que se produzca este hecho. Algunos tipos de apneas son posicionales (se ronca en posición supino), por lo que se podría estimular al paciente cuando ronca para que realizara un cambio de postura a una lateral o prono. Actualmente se realiza con métodos físicos, como por ejemplo, colocando pelotas de tenis cosidas a lo largo de la espalda en una camiseta y poniéndosela para dormir, pero es rechazada por un elevado número de pacientes por ser incómoda. En las apneas posicionales se podría llegar a eliminar totalmente el problema manteniendo la postura adecuada.

Todos los datos recogidos y generados a lo largo de este TFM, se incorporaron a una base de datos relacional diseñada exprofeso. El motor de base de datos elegido es PostgreSQL 9.4. Se decidió crear una base de datos con los siguientes objetivos:

- Almacenamiento de los datos y resultados de una forma estructurada
- Automatización de la entrada de datos y su validación
- Posibilitar la carga de datos desde la futura aplicación que se desarrolle para dispositivos móviles
- Acceso concurrente a los datos por distintos investigadores
- Facilitar el manejo de los datos y su extracción para la realización de análisis
- Facilitar la incorporación de nuevos métodos de análisis de la señal de audio y nuevos atributos o características en el futuro
- Posibilidad de escalado del sistema en el futuro
- Facilitar la copia de seguridad de los datos
- Controlar los privilegios de acceso a los datos, así como disponer de una auditoría de accesos, para cumplir con los requisitos de la LOPD en cuanto a tratamiento de datos de carácter personal.

Se ha implementado la infraestructura necesaria para que todas las herramientas utilizadas en el análisis, transformación y clasificación de los datos obtenidos puedan acceder directamente a la base de datos. Se han desarrollado varios módulos para automatizar los procesos de análisis y transformación de los datos en Matlab®. Para el análisis de los datos se ha utilizado la herramienta WEKA.

Por último, para la validación de los métodos y resultados obtenidos para su aplicación clínica, y la generación automática de patrones de entrenamiento, se han iniciado contactos con diversos hospitales de la Comunidad de Madrid.

ABSTRACT

The Obstructive Sleep Apnea (OSA) is a fairly common sleep disorder, with a prevalence of 2% to 7% of the population. It is estimated that 90% of patients are not diagnosed. The diagnostic test, nocturnal polysomnography (PSG), is expensive and limited by the number of beds available in the Hospital Sleep Unit. OSA has a high social and economic impact due to lack of adequate rest during sleep, thus producing sleepiness states during the vigil that can cause decreased productivity at work and accidents in the handling of machinery and vehicles. Likewise, the situation of hypoxia during sleep can induce various circulatory, neurocognitive and metabolic diseases according to their severity. Therefore, the possibility to have a low cost, simple OSA diagnostic recommendation system, able to reach a large percentage of the population, would be useful for medical services involved in the diagnosis and treatment of OSA.

This Master Thesis proposes a method for analyzing breath sounds during sleep (silent - breathing - snoring), recorded outside of a controlled laboratory environment in order to make a recommendation of OSA diagnostic. Night recordings were conducted by volunteers at their homes, using their own smartphones or a MP3 recorder, without direct supervision. Audio segments containing the desired sounds were extracted, trying that segment contained few background sounds (noise of cars, conversations, bird songs, sounds of mechanisms), whenever possible. These segments were divided in turn in 100 ms windows, which were manually classified into three classes: Silence - Breathing – Snore. Also, for each window, four features were calculated from the extracted audio signal: Zero Crossing Rate, Short Time Energy, autocorrelation coefficient normalized to 1 ms delay, and first predictor coefficient of LPC (Linear Prediction Coding) analysis with 12 poles.

To issue a recommendation for OSA diagnosis, a method for analyzing the audio signal based on the four described audio features has been developed. Different classification methods were evaluated, seeking to offer the best performance in this problem. A RandomForest model applied to features extracted from the audio signal of nocturnal records allowed a correct classification of OSA / No-OSA of 97.33% with a sensitivity of 0.973, specificity of 0.973 and an AUC (area under the curve) of 0.996.

The registered audio features allow online classification of respiratory events, whether Breathing / Silence / Snore, or Non-snore / Snore. In both cases, the classification model with the best performance is a RandomForest. The correct classification of Breathing / Silence / Snore occurs with 85.26% success rate, with a sensitivity of 0.853, specificity of 0.892 and an AUC of 0.944. The classification into Snore / Non-snore occurs with 92.84% success rate, sensitivity of 0.928, specificity of 0.859 and AUC of 0.959.

Additionally, the inclusion of volunteers' biometric data collected through STOP-BANG forms to the classification improved the results. Classification of respiratory events Breathing / Silence / Snore improves to 87.09% of success rate, with sensitivity of 0.871, specificity of 0.903 and AUC of 0.953 by adding the weight, height and age of the volunteer. In the classification of events Non-Snore / Snore, the RandomForest classification improves to a 93.54% success rate, with specificity of 0.935, sensitivity of 0.867 and 0.962 AUC also by adding weight, height and age. Regarding the OSA diagnostic classification on OSA / No-OSA, the RandomForest improves to 97.95% success rate, with sensitivity of 0.98, specificity of 0.979 and 0.997 AUC by adding the gender of the volunteer. Due to the low number of volunteers (7), we were unable to evaluate all collected biometric data, so we will

have to continue studying their contribution in future works, when a larger experimentation population will be available.

Results from diagnosis of OSA allow the future development of a Smartphone application that could be distributed as a screening tool among the population to analyze respiratory sounds and issue a recommendation to visit the doctor if a high probability of suffering OSA is estimated.

The results from the classification of respiratory events open the door to develop systems that deliver stimuli to the user to try to prevent snoring. Some types of apnea are positional (snoring is produced in supine position). It is possible to stimulate the patient when snoring to conduct a change to a lateral or prone posture. Currently this is done with physical methods, such as placing tennis balls sewn on the back in a shirt. However, this is discarded by a large number of patients because it is uncomfortable. Positional apnea can be fully eliminated by maintaining proper posture.

All data collected and generated during this TFM were incorporated into a relational database designed ad-hoc. The chosen database engine is PostgreSQL 9.4. The data base was designed with the following requirements:

- Storage of data and results is done in a structured way
- Automating data entry and validation is automatized
- The system will facilitate the loading of data from the future application that will be developed for mobile devices
- Concurrent access to data by different researchers
- The system facilitates data management and extraction for the analysis
- The system allows the introduction of new methods of analysis of the audio signal and new attributes or features in the future.
- Possibility of scaling the system in the future
- The system allows backup of data
- The system incorporates access control privileges to data and have access audit, to meet the requirements of the Data Protection Act regarding treatment of personal data.

All necessary infrastructure was implemented to allow the direct access of the database by the tools used in the analysis, processing and classification of data. Several modules have been developed in Matlab® to automate the analysis process and data transformation. WEKA was used for data analysis.

Finally, to validate the methods and results for clinical application, we have initiated contacts with various hospitals at Comunidad de Madrid.

ANÁLISIS Y ESTUDIO DE RONQUIDOS PARA LA DETECCIÓN DE TRASTORNOS RESPIRATORIOS DURANTE EL SUEÑO

Autor: D. Francisco Rodríguez González

**Directores: Dr. D. Pablo Varona Martínez
Dr. D. Francisco de Borja Rodríguez Ortiz**

**Grupo de Neurocomputación Biológica
Dpto. de Ingeniería Informática
Escuela Politécnica Superior
Universidad Autónoma de Madrid
Septiembre 2015**

AGRADECIMIENTOS

Espero que este trabajo suponga un paso más en la investigación del diagnóstico automático de SAHS, y que pueda ser de utilidad para futuros investigadores, y por supuesto, para todas aquellas personas que sufren esta enfermedad.

Para mí ha supuesto poder acercarme de nuevo al mundo de la investigación, incorporar nuevos conocimientos y renovadas energías.

En primer lugar, tengo que agradecer la participación de los voluntarios que de forma altruista me han permitido realizar las grabaciones. Sin ellos, este trabajo no habría sido posible. Mi mayor agradecimiento hacia ellos.

Agradecer a la UAM, y al Dpto. de Informática de la EPS, el haberme concedido el acceso y haber podido cursar este Master.

También agradecer sus enseñanzas a los profesores de las asignaturas que he cursado ya que me han abierto la mente a nuevas ideas y posibilidades. Agradecer su apoyo, cada uno con su toque especial. Siempre están dispuestos a ayudar.

A los compañeros de clase, que me han aceptado como uno más, a pesar de que ya peino canas. El tercer tiempo en la cafetería me ayudaba a comprender todos los puntos de vista. Gracias a todos ellos.

Por supuesto, a mis Tutores y Directores de Tesis (Pablo y Paco), por su apoyo y orientación. Admiro su dedicación y paciencia. Muchas gracias de todo corazón.

Y por último agradecer al personal de los distintos Servicios del Hospital Fundación Alcorcón por el tiempo que nos han dedicado y por sus indicaciones de los aspectos clínicos de este trabajo.

El aprendizaje adquirido y los momentos memorables compartidos a lo largo de este Master, hacen que me sienta, ya de por vida, unido a esta Universidad.

Muchas gracias.

Francisco Rodríguez González

ÍNDICE DE CONTENIDOS

1	Introducción	1
1.1	Motivación	1
1.2	Objetivos	2
1.3	Organización de la memoria	3
2	Estado del arte	5
3	Metodología	11
3.1	Obtención de los registros de los voluntarios	14
3.1.1	Reclutamiento de los voluntarios	16
3.1.2	Preparación los formularios	17
3.1.3	Entrega de formularios y sistema de grabación al paciente	18
3.1.4	Contestación a los formularios y realización de grabaciones nocturnas	18
3.1.5	Recogida de los formularios y grabaciones realizadas	19
3.1.6	Introducción en la BBDD de los datos del dispositivo de registro	19
3.1.7	Introducción en la BBDD de los datos del paciente	20
3.1.8	Introducción en la BBDD de los datos de los registros realizados	21
3.1.9	Introducción en la BBDD de los ficheros de audio originales	21
3.2	Normalización y extracción de características	22
3.2.1	Normalización de los ficheros de audio	24
3.2.2	Carga de los datos de audio normalizados en la BBDD	24
3.2.3	Escucha de los ficheros de audio y selección de los segmentos a analizar	25
3.2.4	Extracción de los segmentos a analizar	28
3.2.5	Clasificación manual del segmento de audio	28
3.2.6	Extracción de las características del segmento de audio	29
3.2.6.1	Zero Crossing Rate (Z_k)	30
3.2.6.2	Short Time Energy (E_{sk})	31
3.2.6.3	Coefficiente de autocorrelación normalizado a 1ms de retardo (C_k)	32
3.2.6.4	Primer coeficiente predictor de análisis LPC de 12 polos (L_{pk})	34
3.2.7	Carga en la BBDD de los atributos y clasificación del segmento	36
3.3	Selección del modelo de clasificación	36
3.3.1	Extracción de datos en formato ARFF y MATLAB®	37
3.3.2	Análisis con WEKA y MATLAB®	39
3.4	Operaciones de mantenimiento de los datos	39
3.4.1	Copia de los registros de audio	40
3.4.2	Volcado de la base de datos	41
3.4.3	Archivado en cinta	41
3.5	Entorno Hardware de trabajo	41
4	Resultados	43
4.1	Descripción de los datos recopilados	43
4.2	Clasificación automática de los eventos respiratorios en base a las características del audio	48
4.3	Predicción de pacientes con SAHS usando las características de audio	65
5	Conclusiones	75
6	Trabajo futuro	79
	Referencias	83
	Glosario	89
	Anexos	I
A.	Base de datos de eventos respiratorios	I

B.	Funciones desarrolladas en Matlab® para acceso a la base de datos y cálculo de características	V
C.	Ejemplo de formulario STOP-BANG.....	XI
D.	Descripción de las clases de clasificadores de WEKA y sus parámetros utilizados en este trabajo.....	XIII

ÍNDICE DE FIGURAS

FIGURA 2-1: POLISOMNOGRAFÍA	6
FIGURA 2-2: UN PACIENTE PEDIÁTRICO PREPARADO PARA UNA POLISOMNOGRAFÍA, ST. LOUIS CHILDREN'S HOSPITAL, ST. LOUIS, MISSOURI, 2006. (ROBERT LAWTON).....	7
FIGURA 2-3: SISTEMA DE PSG DOMICILIARIA APNEALINK™ PLUS DE RESMED	7
FIGURA 3-1: DESCRIPCIÓN DE LOS SÍMBOLOS UTILIZADOS EN LOS DIAGRAMAS DESCRIPTIVOS DE LOS PROCESOS	11
FIGURA 3-2: DESCRIPCIÓN GENERAL DEL PROCESO DE ADQUISICIÓN Y TRATAMIENTO DE LOS DATOS	13
FIGURA 3-3: DESCRIPCIÓN DEL SUBPROCESO DE OBTENCIÓN DE REGISTROS DE LOS VOLUNTARIOS	15
FIGURA 3-4: MICRÓFONO RØDE SMARTLAV.....	19
FIGURA 3-5: DESCRIPCIÓN DEL SUBPROCESO DE NORMALIZACIÓN Y EXTRACCIÓN DE CARACTERÍSTICAS.....	23
FIGURA 3-6: CAPTURA DE PANTALLA DE LA APLICACIÓN AUDACITY® MOSTRANDO LA FORMA DE ONDA Y SU ESPECTRO DE UN REGISTRO DE RONQUIDOS.	26
FIGURA 3-7: REPRESENTACIÓN DEL ZK CALCULADO, FORMA DE ONDA Y CLASE ASIGNADA PARA UNA PEQUEÑA PORCIÓN DE UN SEGMENTO DE AUDIO	31
FIGURA 3-8: REPRESENTACIÓN DE ESK CALCULADO, FORMA DE ONDA Y CLASE ASIGNADA PARA UNA PEQUEÑA PORCIÓN DE UN SEGMENTO DE AUDIO	32
FIGURA 3-9: REPRESENTACIÓN DE Ck CALCULADO, FORMA DE ONDA Y CLASE ASIGNADA PARA UNA PEQUEÑA PORCIÓN DE UN SEGMENTO DE AUDIO	33
FIGURA 3-10: REPRESENTACIÓN DE Ck CALCULADO, FORMA DE ONDA Y CLASE ASIGNADA EN OTRO SEGMENTO DE AUDIO	33
FIGURA 3-11: MODELO DE TUBOS RESONANTES DEL TRACTO VOCAL	34
FIGURA 3-12: REPRESENTACIÓN DE LPK CALCULADO, FORMA DE ONDA Y CLASE ASIGNADA PARA UNA PEQUEÑA PORCIÓN DE UN SEGMENTO DE AUDIO	35
FIGURA 3-13: DESCRIPCIÓN DEL SUBPROCESO DE SELECCIÓN DEL MODELO DE CLASIFICACIÓN....	37
FIGURA 3-14: MATLAB® DATABASE EXPLORER	38
FIGURA 3-15: ACCESO A LA BASE DE DATOS DESDE WEKA.....	39
FIGURA 3-16: DESCRIPCIÓN DEL SUBPROCESO DE OPERACIONES DE MANTENIMIENTO DE LOS DATOS	40
FIGURA 3-17: ESQUEMA DEL ENTORNO HARDWARE DE TRABAJO	42
FIGURA 4-1: DISTRIBUCIÓN DE LOS EVENTOS RESPIRATORIOS CLASIFICADOS	44
FIGURA 4-2: DISTRIBUCIÓN DE CLASES POR MODELO DE GRABADORA	45
FIGURA 4-3: DISTRIBUCIÓN DE CLASES POR PACIENTE	45
FIGURA 4-4: DISTRIBUCIÓN DE CLASES POR GÉNERO.....	46
FIGURA 4-5: DISTRIBUCIÓN DE CLASES POR GRUPO DE EDAD	47
FIGURA 4-6: DISTRIBUCIÓN DE CLASES POR SAHS DIAGNOSTICADO	47
FIGURA 4-7: DISTRIBUCIÓN DE CLASES POR GRUPO DE BMI.....	48
FIGURA 4-8: DISTRIBUCIÓN DE ZK EN FUNCIÓN DE LA CLASE DE EVENTO RESPIRATORIO.....	50
FIGURA 4-9: DISTRIBUCIÓN DE ESKEN FUNCIÓN DE LA CLASE DE EVENTO RESPIRATORIO.....	50
FIGURA 4-10: DISTRIBUCIÓN DE Ck EN FUNCIÓN DE LA CLASE DE EVENTO RESPIRATORIO	51
FIGURA 4-11: DISTRIBUCIÓN DE LPK EN FUNCIÓN DE LA CLASE DE EVENTO RESPIRATORIO.....	52
FIGURA 4-12: RELACIONES ENTRE LOS ATRIBUTOS DE FK PARA LOS EVENTOS RESPIRATORIOS	53
FIGURA 4-13: RED NEURONAL DEL PERCEPTRÓN DE UNA CAPA OCULTA CON TRES NEURONAS	57
FIGURA 4-14: DISTRIBUCIÓN DE LA ALTURA DE LOS VOLUNTARIOS POR CLASE DE EVENTO RESPIRATORIO	58
FIGURA 4-15: DISTRIBUCIÓN DEL PESO DE LOS VOLUNTARIOS POR CLASE DE EVENTO RESPIRATORIO	59
FIGURA 4-16: DISTRIBUCIÓN DE LA EDAD DE LOS VOLUNTARIOS POR CLASE DE EVENTO RESPIRATORIO	59

FIGURA 4-17: DISTRIBUCIÓN DE BMI DE LOS VOLUNTARIOS POR CLASE DE EVENTO RESPIRATORIO	60
FIGURA 4-18: DISTRIBUCIÓN DE GENERO DE LOS VOLUNTARIOS POR CLASE DE EVENTO RESPIRATORIO	60
FIGURA 4-19: DISTRIBUCIÓN DE ZK POR DIAGNÓSTICO DE SAHS	65
FIGURA 4-21: DISTRIBUCIÓN DE CK POR DIAGNÓSTICO DE SAHS	66
FIGURA 4-22: DISTRIBUCIÓN DE LPK POR DIAGNÓSTICO DE SAHS.....	67
FIGURA 4-23: DISTRIBUCIÓN DE ALTURA DE LOS VOLUNTARIOS POR DIAGNÓSTICO DE SAHS	69
FIGURA 4-24: DISTRIBUCIÓN DEL PESO DE LOS VOLUNTARIOS POR DIAGNÓSTICO DE SAHS	70
FIGURA 4-25: DISTRIBUCIÓN DE LA EDAD DE LOS VOLUNTARIOS POR DIAGNÓSTICO DE SAHS	70
FIGURA 4-26: DISTRIBUCIÓN DEL BMI DE LOS VOLUNTARIOS POR DIAGNÓSTICO DE SAHS.....	71
FIGURA 4-27 DISTRIBUCIÓN DEL GÉNERO DE LOS VOLUNTARIOS POR DIAGNÓSTICO DE SAHS	71
FIGURA A-1: DIAGRAMA DE LA BASE DE DATOS DE RONQUIDOS (v2.8).....	II
FIGURA C-1: EJEMPLO DE FORMULARIO STOP-BANG	XI

ÍNDICE DE TABLAS

TABLA 3-1: DISPOSITIVOS DE GRABACIÓN UTILIZADOS Y SUS CARACTERÍSTICAS	20
TABLA 4-1: DATOS BIOMÉTRICOS Y NÚMERO DE REGISTROS DE LOS VOLUNTARIOS	43
TABLA 4-2: DISTRIBUCIÓN DE LOS EVENTOS RESPIRATORIOS CLASIFICADOS	43
TABLA 4-3: DISTRIBUCIÓN DE CLASES POR GRABADORA.....	44
TABLA 4-4: DISTRIBUCIÓN DE CLASES POR PACIENTE	45
TABLA 4-5: DISTRIBUCIÓN DE CLASES POR GÉNERO	46
TABLA 4-6: DISTRIBUCIÓN DE CLASES POR EDAD	46
TABLA 4-7: DISTRIBUCIÓN DE CLASES POR SAHS DIAGNOSTICADO	47
TABLA 4-8: DISTRIBUCIÓN DE CLASES POR GRUPO DE BMI.....	48
TABLA 4-9: MATRIZ DE CORRELACIÓN DE LOS ATRIBUTOS DE FK	54
TABLA 4-10: RESULTADOS DE LOS DISTINTOS CLASIFICADORES PARA LOS EVENTOS RESPIRATORIOS	56
TABLA 4-11: CLASIFICACIÓN DE EVENTOS RESPIRATORIOS CON RANDOMFOREST AÑADIENDO ATRIBUTOS STOP-BANG	61
TABLA 4-12: CLASIFICACIÓN DE EVENTOS RESPIRATORIOS CON J48 AÑADIENDO ATRIBUTOS STOP-BANG	62
TABLA 4-13: RESULTADOS DE LOS DISTINTOS CLASIFICADORES PARA LOS EVENTOS RONQUIDO/NO_RONQUIDO	63
TABLA 4-14: CLASIFICACIÓN DE RONQUIDO/NO_RONQUIDO CON RANDOMFOREST MÁS STOP-BANG	64
TABLA 4-15: RESULTADOS DE LOS DISTINTOS CLASIFICADORES PARA LA PREDICCIÓN DE SAHS..	68
TABLA 4-16: COMPARACIÓN DE CLASIFICADORES PARA LA PREDICCIÓN DE SAHS	68
TABLA 4-17: RESULTADOS DE LOS DISTINTOS CLASIFICADORES PARA LA PREDICCIÓN DE SAHS INCLUYENDO EL GÉNERO	72
TABLA 4-19: RESULTADOS DE LOS DISTINTOS CLASIFICADORES PARA LA PREDICCIÓN DE SAHS INCLUYENDO EL PESO.....	73
TABLA 4-20: COMPARACIÓN DE CLASIFICADORES PARA LA PREDICCIÓN DE SAHS USANDO EL GÉNERO DEL VOLUNTARIO	74
TABLA 4-21: COMPARACIÓN DE LOS CLASIFICADORES EN LA PREDICCIÓN DE SAHS SIN Y CON ATRIBUTO <i>GÉNERO DEL VOLUNTARIO</i>	74

1 Introducción

1.1 Motivación

El Síndrome de Apnea-Hipoapnea del Sueño (SAHS) es un trastorno del sueño bastante común relacionado con la alteración de la respiración mientras se duerme. Consiste en la oclusión total o parcial de la vía aérea superior que produce una disminución de la concentración de oxígeno en sangre, lo cual puede inducir diversas enfermedades circulatorias, neurocognitivas y metabólicas en función de su severidad.

Es un síndrome con un alto impacto social y económico, ya que una de sus características es la falta de un descanso adecuado durante el sueño, por lo que se producen estados de somnolencia durante la vigilia que pueden afectar a la realización de tareas donde se requiera una concentración adecuada, por lo que aumenta el riesgo de sufrir accidentes por falta de atención. En la legislación española, se recoge padecer SAHS como causa suficiente para denegar el permiso de conducir, debido al riesgo de somnolencia al conducir vehículos y maquinaria pesada. Así mismo, las personas que padecen SAHS presentan un mayor porcentaje de absentismo laboral de larga duración, menores niveles de productividad y un nivel alto de distrés¹ psicológico.

Se estima que la prevalencia del SAHS en el mundo es del 2% al 7% de la población, dependiendo de la raza, género y lugar de residencia. Así mismo, se estima que el 90% de los enfermos de SAHS no están diagnosticados ni están en tratamiento.

El SAHS se diagnostica mediante una Polisomnigrafía (PSG) mientras el paciente duerme durante la noche. Es una prueba cara y limitada por el número de camas disponibles en la Unidad del Sueño del hospital. Aunque se han desarrollado técnicas de PSG domiciliaria, sigue siendo un factor limitante el número de dispositivos disponibles. El poder disponer de un sistema de recomendación de diagnóstico de bajo coste, sencillo y que alcance a un gran porcentaje de la población sería de gran utilidad para los Servicios Médicos implicados en el diagnóstico y tratamiento del SAHS.

La motivación del presente Trabajo Fin de Master es encontrar un método de recomendación de diagnóstico de SAHS en base a un modelo de extracción de características y clasificación de los sonidos producidos por la respiración (silencio – respiración – ronquido) durante el sueño, que pueda ser utilizado en un dispositivo portátil, ya sea un “*Smartphone*” de gran difusión, un sistema dedicado, o una combinación de ambos.

Así mismo, se propone analizar si la incorporación de distintos datos biométricos adicionales y fáciles de obtener mediante cuestionarios (peso, altura, edad, sexo, etc.), ayudaría en la clasificación de los eventos respiratorios y en el diagnóstico del SAHS.

El uso de este sistema permitiría realizar un cribado de un porcentaje alto de la población a un coste asumible, de manera que se revelen posibles pacientes de SAHS no detectados para

¹ Actualmente se tiende a distinguir entre el estrés positivo, denominado “eustrés”, y el negativo, denominado “distrés”.

su posterior diagnóstico clínico. Así mismo, serviría para el seguimiento de la evolución de los pacientes ya tratados o en tratamiento.

El desarrollo de esta aplicación de análisis de posible diagnóstico de SAHS en dispositivos “*Smartphone*” permitiría una obtención masiva de datos e incorporarlos en un repositorio central, lo que puede permitir desarrollar nuevas técnicas de investigación y diagnóstico dentro del contexto del “*big data*” biomédico.

1.2 Objetivos

Partiendo de la premisa de que los modelos matemáticos a desarrollar para la posible detección de SAHS y eventos espiratorios se han de incorporar en un dispositivo portátil, se definen los siguientes objetivos:

1. Encontrar y definir las características de audio que mejor definan la separación de los tres eventos respiratorios Ronquido – Respiración – Silencio. Los métodos a desarrollar han de cumplir:
 - a. Tener bajo coste computacional, tanto a nivel de CPU como de memoria utilizada.
 - b. Ser lo suficientemente robustos para ser efectivos con bajas frecuencias de muestreo.
 - c. Ser lo suficientemente robustos para poder ser utilizados en distintos dispositivos con variaciones en la calidad de sus sistemas de adquisición de audio.
 - d. Ser lo suficientemente robustos para no verse influidos por la variabilidad de paciente (edad, sexo, etc.).
 - e. Poder realizar la extracción de características *on-line*.

La clasificación *on-line* se persigue para en un desarrollo futuro poder detectar eventos SAHS en el momento en el que se producen, e intentar aplicar al paciente estímulos en ese mismo instante, que corrijan o palien la causa de la apnea.

2. Realizar una predicción de si el paciente que se está registrando es susceptible de sufrir SAHS, para emitir una recomendación de visita al especialista médico. Esta predicción se puede realizar *off-line*.
3. Seleccionar los datos biométricos adicionales obtenidos mediante formularios, que ayuden a mejorar la clasificación de las características de audio y la predicción de la posibilidad de sufrir SAHS.
4. Consolidar toda la información adquirida, datos y resultados, en una base de datos que conforme los cimientos del sistema de recogida de datos en un futuro despliegue de una aplicación de “*Smartphone*”. Esta base de datos ha de ayudar a estructurar toda esta información para su uso en futuras investigaciones.
5. Contactar con los equipos médicos relacionados con el diagnóstico y tratamiento de SAHS, para validar la aplicación clínica de los métodos y procedimientos desarrollados.

1.3 Organización de la memoria

La memoria consta de los siguientes capítulos:

2. **Estado del arte:** se presenta el estado del arte relativo a la detección y diagnóstico de SAHS, técnicas clínicas utilizadas, detección de eventos (sonidos respiratorios) durante el sueño y evolución hacia dispositivos de diagnóstico móviles (basados en *smartphones*).
 3. **Metodología:** se describe la metodología seguida para la localización de los voluntarios, la obtención de los registros de audio y datos experimentales, creación de una base de datos centralizada para la estructuración de los datos, tratamiento de los ficheros de audio para su normalización, carga de los ficheros de audio en la base de datos, extracción de características de los ficheros de audio, etiquetado manual de los eventos de audio, carga de todos los datos obtenidos en la base de datos y las operaciones de mantenimiento de los datos.
 4. **Resultados:** se describen los datos experimentales recopilados, los resultados de la detección de los eventos/sonidos respiratorios, los resultados de diagnóstico de SAHS, y las mejoras obtenidas al incorporar datos de formularios STOP-BANG.
 5. **Conclusiones:** se recogen las principales conclusiones extraídas de este Trabajo Fin de Master, y se describe el grado de cumplimiento de los objetivos de proyecto del TFM.
 6. **Trabajo futuro:** Se describen las futuras líneas de trabajo en la mejora de los métodos y técnicas utilizados, así como nuevas investigaciones que desemboquen en la realización de una Tesis Doctoral.
- A. **Anexos:** Adicionalmente, se incluyen en la parte final una serie de anexos que complementan la información presentada en este documento. Principalmente, se constituye como información complementaria a la ya expuesta en el documento principal y que desarrolla aspectos cuya inclusión detallada en dicho manuscrito podría desviar la atención de las verdaderas aportaciones del trabajo principal.

2 Estado del arte

Los trastornos del sueño son bastante comunes y contribuyen a un gran número de problemas de salud (Idzikowski, 2012; Judd, 2010), incluyendo enfermedades cardiovasculares (Furukawa et al., 2010; Kasai & Bradley, 2011), mentales (Sanchez, Bermudez, & Buela-Casal, 2003) y metabólicas (Jurado Gámez et al., 2010). Estas alteraciones se pueden incluso presentar a edades tempranas (Vlahandonis et al., 2014).

Uno de los trastornos más comunes es el relacionado con alteraciones de la respiración durante el sueño. La oclusión de la vía aérea superior puede ser parcial o total (síndrome de la apnea del sueño, OSA – *Obstructive Sleep Apnea* en inglés, SAHS – *Síndrome de Apnea Hipoapnea del Sueño* en español) (Punjabi, 2008). Se define un evento SAHS como la ausencia de flujo de aire durante al menos 10 segundos seguidos. Los distintos grados de severidad de SAHS se definen en función del número de eventos y su duración total respecto al total del tiempo de sueño (American Academy of Sleep Medicine, 2001). Esta falta de oxígeno durante el sueño puede producir enfermedades como la isquemia cardíaca, el ictus cerebral, la diabetes tipo II, disfunciones neurocognitivas (Pack, 2012), y un aumento del riesgo de sufrir accidentes por falta de concentración (somnolencia) (Lloberes et al., 2000).

Se estima que en Australia este síndrome tiene un impacto en costes equivalente al 0.8% de su PIB en 2006, ya sea por bajas laborales, costes médicos o accidentes de tráfico (Behar, Roebuck, Domingos, Gederi, & Clifford, 2013). La prevalencia de SAHS en el mundo es del 2% al 7.5% de la población, dependiendo de la raza, género y lugar de residencia.

En España, el Ministerio del Interior está preocupado por cómo afecta el SAHS a las aptitudes físicas y psicológicas necesarias para conducir vehículos, tal y como recoge la *Línea de Investigación 2.D: Efectos de Determinadas Enfermedades Cónicas* dentro del Plan de Investigación en Seguridad Vial y Movilidad 2013-2016 de la Dirección General de Tráfico (Dirección General de Tráfico, 2014).

Así mismo, la legislación española ya incluye el SAHS diagnosticado como patología incapacitante para manejar vehículos, tal y como recoge el R.D. 818/2009, de 8 de mayo, por el que se aprueba el Reglamento General de Conductores. En su *Anexo IV, Aptitudes psicofísicas requeridas para obtener o prorrogar la vigencia del permiso o de la licencia de conducción*, en el apartado 7.2 *Trastornos del sueño*, se indica:

No se permiten el síndrome de apneas obstructivas del sueño, los trastornos relacionados con éste, ni otras causas de excesiva somnolencia diurna (Ministerio del Interior, 2009)

Un reciente estudio realizado en España (Jurado-Gámez, Guglielmi, Gude, & Buela-Casal, 2015), concluye que el SAHS es una enfermedad con repercusiones importantes tanto económicas como sociales. Los pacientes con SAHS, desde el punto de vista laboral, presentan un mayor porcentaje de absentismo de larga duración (más de 30 días), menores niveles de productividad y un nivel alto de distrés² psicológico.

² Actualmente se tiende a distinguir entre el estrés positivo, denominado “eustrés”, y el negativo, denominado “distrés”.

El diagnóstico de SAHS se realiza mediante la revisión de la historia clínica y el examen físico del paciente. La prueba definitiva se obtiene en el laboratorio del sueño mediante una polisomnografía (PSG) clínica. Durante la sesión de PSG se registra y monitoriza electroencefalograma (EEG), electrooculograma (EOG), electromiograma (EMG) del mentón, movimientos de las extremidades, electrocardiograma (ECG), sonidos de la vía aérea superior, respiración (flujo de aire y esfuerzo respiratorio), saturación de oxígeno en sangre (SpO₂), capnografía (nivel de CO₂ en el aire expulsado), pletismografía (movimientos de la caja torácica y del abdomen), posición del cuerpo y registros en video para la observación del comportamiento del paciente (American Association of Sleep Technologist, 2012). El problema es que este método de diagnóstico es caro, superior a los 1000€ por sesión/paciente (Behar, Roebuck, Domingos, et al., 2013) debido a que requiere la monitorización del paciente durante una noche completa, por lo que su aplicación está limitada por el número de camas disponibles en el laboratorio del sueño y el número de especialistas en el turno; las listas de espera pueden superar los dos años (A Roebuck et al., 2014). A este problema hay que añadir que la realización del PSG es muy incómoda para el paciente. En la Figura 2-1 podemos observar algunos de los electrodos y sensores que se conectan al paciente para realizar la prueba. Una complicación adicional es el realizar esta prueba a niños, pues aunque se intenta que el entorno sea lo menos agresivo para el menor, sigue siendo un entorno desconocido y traumático. En la Figura 2-2 se puede observar a un niño preparado para una polisomnografía.



Figura 2-1: Polisomnografía³

Para reducir los costes financieros de la diagnosis de SAHS y llegar a un mayor número de pacientes, se desarrollaron sistemas de PSG alternativos, con un menor número de sensores, para ser utilizados por el paciente en su propia casa sin necesidad de acudir al hospital. El problema de los PSG domiciliarios es que el paciente ha de colocarse él mismo los electrodos y sensores, por lo que muchas veces las medidas obtenidas no son válidas (Masa et al., 2013). En la Figura 2-3 podemos observar un sistema comercial de PSG domiciliaria.

Para la realización de la PSG domiciliaria, cada Hospital o Unidad del Sueño tiene sus propios procedimientos. Por ejemplo, en el Hospital El Escorial envían un técnico a casa del paciente con el equipo y le explican cómo se lo ha de colocar. Por la noche, el paciente se

³ Fuente: <https://sleep.wustl.edu>

coloca él mismo los diferentes electrodos y el sistema de registro. A la mañana siguiente, el técnico vuelve para a recoger el equipo.



Figura 2-2: Un paciente pediátrico preparado para una polisomnografía, St. Louis Children's Hospital, St. Louis, Missouri, 2006. (Robert Lawton)⁴

Otra solución es que el paciente acuda al centro hospitalario a que un técnico le coloque los electrodos y el sistema de registro, y regrese a dormir a su casa. A la mañana siguiente vuelve al hospital a que le retiren el dispositivo de medida. Esta es el procedimiento que utilizan en el Hospital Universitario Fundación Alcorcón. En cualquier caso, la prueba de PSG domiciliaria se realiza sin la supervisión de un técnico, pues aún en el caso de que acuda al hospital a que le coloquen el sistema de registro, no se tiene control de lo que pueda hacer el paciente en su domicilio.



Figura 2-3: Sistema de PSG domiciliaria ApneaLink™ Plus de ResMed⁵

Se estima que 90% de los enfermos de SAHS no están diagnosticados, ni están en tratamiento. Por lo tanto, sería de gran ayuda disponer de un método de diagnóstico sencillo, de bajo coste y que alcance a un mayor porcentaje de la población (Behar, Roebuck, Shahid, et al., 2013).

⁴ Fuente: Wikipedia (licencia Creative Commons Genérica de Atribución/Compartir-Igual 2.5)

⁵ Fuente: http://www.resmed.com/la/products/apnealink_plus/apnealink-plus.html?nc=patients

Una manifestación de la obstrucción parcial de la vía aérea superior es la vibración de sus tejidos blandos al ocluirse. Produce el sonido característico que conocemos como ‘ronquido’. Los pacientes que sufren de SAHS son roncadores; el roncar no indica directamente tener SAHS, pero puede ser un indicio de que se está evolucionando hacia el síndrome. Un método sencillo para detectar pacientes con alteraciones de la respiración durante el sueño es el estudio de los sonidos producidos (silencios, respiración, ronquido). El uso de estos métodos permitiría aplicar terapias preventivas antes de que se llegase el estadio de SAHS (A Roebuck et al., 2014).

Hoy día, el avance de la tecnología en dispositivos móviles (Smartphone, tabletas, etc...), tanto a nivel computacional como de comunicaciones, así como la reducción de su coste de adquisición, ofrece nuevas oportunidades de innovación en las áreas relacionadas con la salud. Recientemente, Behar et al. (Behar, Roebuck, Domingos, et al., 2013) realizaron una revisión de las aplicaciones de Smartphone disponibles para la monitorización del sueño. Estas aplicaciones pueden hacer uso de los sensores de los teléfonos (micrófonos, acelerómetros, accesorios externos, etc.), y/o realizar los cuestionarios STOP-BANG (Frances Chung et al., 2008) utilizados por los anestesistas. El cuestionario STOP-BANG es un formulario de ocho preguntas que predicen el SAHS con una sensibilidad del 78-90% y una especificidad del 85-90% dependiendo de la severidad de la enfermedad (Frances Chung, Yang, & Liao, 2013). Aunque existe un alto número de aplicaciones, los autores concluyen que ninguna se basa en una evidencia científica fundamentada, con la excepción de las que simplemente implementan el cuestionario STOP-BANG.

En la bibliografía podemos encontrar multitud de métodos y procedimientos para la detección y separación de sonidos relacionados con la respiración durante el sueño (silencio-respiración-ronquido):

- Energía de la señal y Zero Crossing Rate (Abeyratne, Wakwella, & Hukins, 2005)
- análisis espectrales (Sola-Soler, Jane, Fiz, & Morera, 2003)
- entropía multiescala (Aoife Roebuck & Clifford, 2012)
- modelos de Markov ocultos (Snider & Kain, 2013)
- análisis de modificaciones en la voz (Goldshtein, Tarasiuk, & Zigel, 2011; Montero Benavides et al., 2014, 2015; Solé-Casals et al., 2014)
- análisis con Wavelets (Matsiki, Deligianni, Vlachogianni-Daskalopoulou, & Hadjileontiadis, 2007)
- análisis mediante Linear Predictive Coding (LPC) y coeficientes cepstrales (CC y MFCC) (Snider & Kain, 2013)
- análisis del intervalo de tiempo entre ronquidos durante toda la noche (Mesquita, Sola-Soler, Fiz, Morera, & Jane, 2012)
- intensidad del ronquido (Maimon & Hanly, 2010)
- combinaciones de distintos métodos (Karunajeewa, Abeyratne, & Hukins, 2008).

Muchos de los métodos descritos no son de aplicación en un sistema *on-line*, ya que necesitan del procesado de segmentos de audio de bastante duración o una potencia de cálculo no alcanzable por un dispositivo móvil. Por ejemplo, en el método de entropía multiescala, las escalas que mejores resultados proporcionaban eran las de 6, 21 y 30 segundos con una especificidad del 90.5% y un valor predictivo positivo de 83.5% (A Roebuck et al., 2014).

Siguiendo con el ejemplo anterior, estas escalas de tiempo de 6, 21 o 30 segundos son demasiado grandes si queremos tener un sistema de detección *on-line* que nos permita actuar sobre el paciente para intentar corregir la causa del ronquido/apnea mediante un sistema que se desarrollará en el futuro. Este sistema intentará corregir la causa de la apnea, por ejemplo, las debidas a la postura (Oksenberg & Silverberg, 1998), forzando al paciente a adoptar una postura lateral en la que el colapso de la vía aérea sea menor (Dan Anker Hofsoy, Clauss, & Wolf, 2009; van Maanen et al., 2013). En la actualidad, el método para forzar la posición lateral es mediante la colocación de pelotas de tenis cosidas a una camiseta a lo largo de la espalda, tal y como nos explicaron en el Servicio de Neumología del Hospital Fundación Alcorcón (Counter & Wilson, 2004). También sería de aplicación en roncadores que no sufren SAHS como sistema de prevención de la apnea (Benoist, Morong, van Maanen, Hilgevoord, & de Vries, 2014; D A Hofsoy, Clauss, & Wolf, 2009).

3 Metodología

En esta sección se describe el proceso de obtención de datos de los voluntarios, así como su almacenamiento y procesamiento para la extracción de características y clasificación de eventos.

Los símbolos utilizados en los diagramas descriptivos de los distintos procesos de recogida y análisis de los datos son los siguientes:


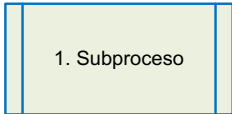
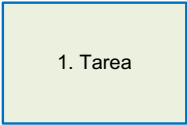
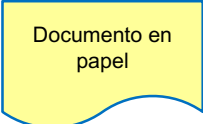

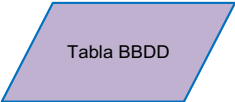

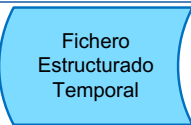
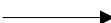
SÍMBOLO	DESCRIPCIÓN
	<p>Diagrama de Proceso. Describe las tareas y subprocesos a ejecutar, así como el acceso a datos y herramientas.</p> <ul style="list-style-type: none"> Barra horizontal azul: Identificador del proceso Barra vertical azul: indica subsecciones del proceso (opcional) Columna amarilla: Indica datos y herramientas utilizadas por la tarea Columna verde: indica las tareas/subprocesos a realizar por una Función: persona, grupo, Unidad Funcional, etc.
	<p>Subproceso definido en otro diagrama. El texto indica el identificador del subproceso. Solo aparece en las columnas de Función.</p>
	<p>Tarea a realizar por una función. El texto indica el identificador de la tarea Solo aparece en las columnas de Función.</p>
	<p>Documento en papel generado o utilizado como entrada por una tarea o subproceso. Solo aparece en las columnas de datos/herramientas</p>
	<p>Fichero secuencial generado o utilizado como entrada por una tarea o subproceso. Solo aparece en las columnas de datos/herramientas</p>
	<p>Tabla de la base de datos accedida por una tarea o subproceso. El texto indica el nombre de la tabla. Solo aparece en las columnas de datos/herramientas</p>
	<p>Fichero secuencial temporal generado o utilizado como entrada por una tarea o subproceso. Solo aparece en las columnas de datos/herramientas</p>
	<p>Fichero estructurado temporal generado o utilizado como entrada por una tarea o subproceso. Solo aparece en las columnas de datos/herramientas</p>
	<p>Flujo. Puede ser de datos o de ejecución de las tareas del proceso.</p>

Figura 3-1: Descripción de los símbolos utilizados en los diagramas descriptivos de los procesos

El proceso general para la recogida de datos, extracción de características de audio, clasificación de los eventos respiratorios y diagnóstico de SAHS consta de cuatro subprocesos:

1. Obtención de los registros de los voluntarios
2. Normalización y extracción de características
3. Selección del modelo de clasificación
4. Operaciones de mantenimiento de los datos

En la Figura 3-2 se muestra el flujo de trabajo y las entradas y salidas de los cuatro subprocesos.

Cada subproceso se compone de tareas concretas para alcanzar la funcionalidad requerida. En las siguientes subsecciones se detallan las tareas que componen cada subproceso.

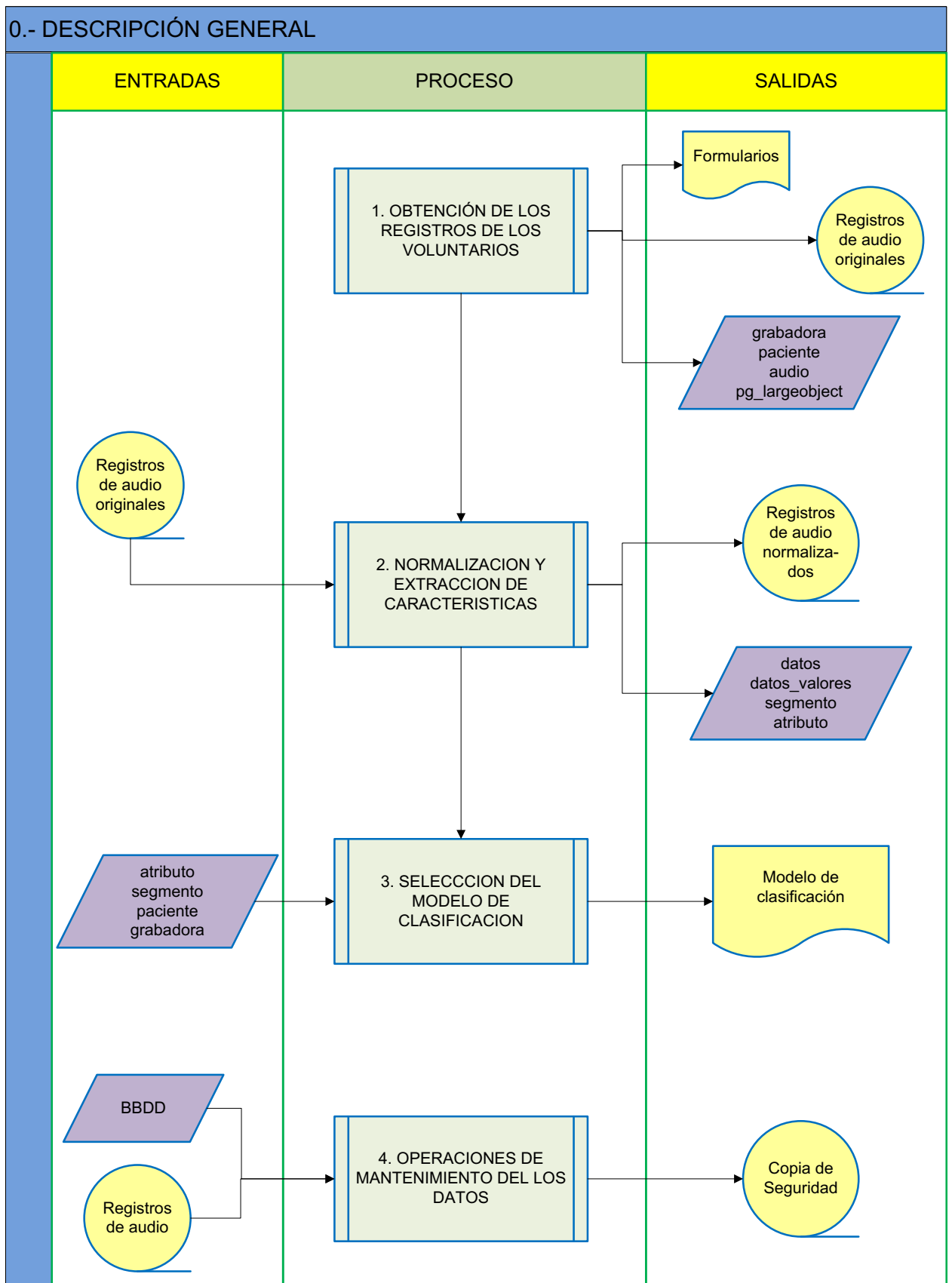


Figura 3-2: Descripción general del proceso de adquisición y tratamiento de los datos

3.1 Obtención de los registros de los voluntarios

En este subproceso se realiza la recogida de datos de los voluntarios y su tabulación en una base de datos diseñada ex profeso para este trabajo.

Se decidió crear una base de datos con los siguientes objetivos:

- Almacenamiento de los datos y resultados de una forma estructurada
- Automatización de la entrada de datos y su validación
- Posibilitar la carga de datos desde la futura aplicación que se desarrolle para dispositivos móviles
- Acceso concurrente a los datos por distintos investigadores
- Facilitar el manejo de los datos y su extracción para la realización de análisis
- Facilitar la incorporación de nuevos métodos de análisis de la señal de audio y nuevos atributos o características en el futuro.
- Posibilidad de escalado del sistema en el futuro
- Facilitar la copia de seguridad de los datos
- Controlar los privilegios de acceso a los datos, así como disponer de una auditoría de accesos, para cumplir con los requisitos de la LOPD en cuanto a tratamiento de datos de carácter personal.

Esta base de datos está construida sobre PostgreSQL 9.4⁶. En el anexo A se recoge el esquema y características de la base de datos diseñada.

Todos los datos recogidos, son introducidos en la base de datos para su posterior tratamiento y análisis:

- Datos del voluntario
- Características de los equipos de grabación
- Registros de audio
- Atributos y características del análisis de audio
- Clasificación de los eventos respiratorios

Para la transformación y análisis de los registros de audio se ha utilizado Matlab®⁷ (ver sección 3.2) y se han desarrollado varios módulos para automatizar los procesos (ver anexo A). La clasificación de los eventos respiratorios se ha realizado con WEKA⁸ (ver sección 3.3).

En la Figura 3-3 se muestra el flujo de trabajo y datos para la realización del proceso de obtención de los registros de los voluntarios.

A continuación se detallan las tareas que componen este subproceso.

⁶ PostgreSQL: <http://www.postgresql.org>

⁷ <http://es.mathworks.com/products/matlab/?refresh=true>

⁸ <http://www.cs.waikato.ac.nz/~ml/weka/>

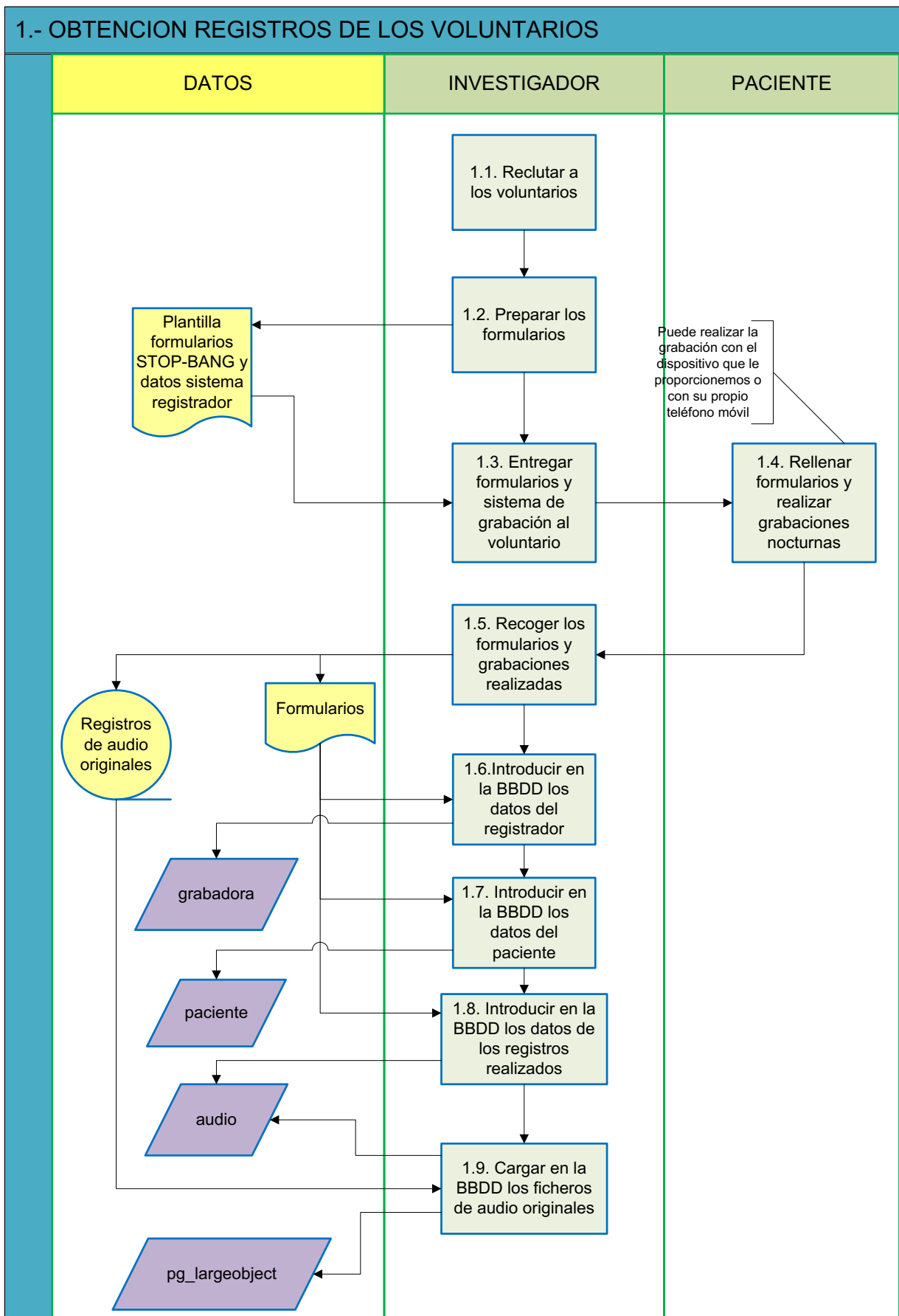


Figura 3-3: Descripción del subproceso de obtención de registros de los voluntarios

3.1.1 Reclutamiento de los voluntarios

Dada la naturaleza del proyecto, se ha necesitado de la colaboración de voluntarios que se prestaran a ser grabados durante la noche.

A la fecha de presentación de este trabajo, se han prestado a colaborar siete (7) personas:

- 5 varones y 2 mujeres
- 5 roncan de forma clara
- 2 con diagnóstico clínico positivo de SAHS mediante PSG
- 1 con diagnóstico clínico negativo de SAHS mediante PSG
- la edad oscila entre los 6 y los 57 años

Ya que en el presente trabajo se están manejando datos que pueden ser considerados de carácter personal, hay que remitirse a la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal (Jefatura del Estado, 1999); y al R.D. 1720/2007, de 21 de diciembre, por el que se aprueba el Reglamento de desarrollo de la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal (Justicia & Ministerio De Justicia, 2012).

El R.D. 1720/2007 comparte con la Ley Orgánica la finalidad de hacer frente a los riesgos que para los derechos de la personalidad pueden suponer el acopio y tratamiento de datos personales. Es de aplicación a los datos de carácter personal registrados en soporte físico, que los haga susceptibles de tratamiento, y a toda modalidad de uso posterior de estos datos por los sectores público y privado. En su Artículo 5, define:

“f) Datos de carácter personal: Cualquier información numérica, alfabética, gráfica, fotográfica, acústica o de cualquier otro tipo concerniente a personas físicas identificadas o identificables.”

“g) Datos de carácter personal relacionados con la salud: las informaciones concernientes a la salud pasada, presente y futura, física o mental, de un individuo. En particular, se consideran datos relacionados con la salud de las personas los referidos a su porcentaje de discapacidad y a su información genética.”

“k) Fichero: Todo conjunto organizado de datos de carácter personal, que permita el acceso a los datos con arreglo a criterios determinados, cualquiera que fuere la forma o modalidad de su creación, almacenamiento, organización y acceso.”

“o) Persona identificable: toda persona cuya identidad pueda determinarse, directa o indirectamente, mediante cualquier información referida a su identidad física, fisiológica, psíquica, económica, cultural o social. Una persona física no se considerará identificable si dicha identificación requiere plazos o actividades desproporcionados.”

Todos los datos recogidos se han tratado de forma anónima, de manera que los registros obtenidos no se puedan relacionar los datos biométricos y de salud con ninguna persona física, y así cumplir con la L.O. 15/1999 y el R.D. 1720/2007, de manera que no es necesario inscribir el fichero obtenido en el Registro de la AEPD.

3.1.2 Preparación los formularios

En clínica, los cuestionarios se utilizan normalmente como una primera criba sencilla para detectar trastornos del sueño. Principalmente se han desarrollado en el ámbito de la anestesiología, ya que los pacientes con SAHS no diagnosticada pueden tener complicaciones graves en el post-operatorio (A Roebuck et al., 2014).

Distintos grupos investigadores han desarrollado formularios más o menos extensos: Berlín (Kang et al., 2013), Karolinska (Westerlund, Brandt, Harlid, Åkerstedt, & Trolle Lagerros, 2014), etc. Abrishami, Khajehdehi, & Chung (2010) realizaron una revisión de los principales cuestionarios utilizados en la predicción del SAHS.

Chug et al. (2008) desarrollaron el formulario STOP, mejorándolo posteriormente con el modelo STOP-BANG (Frances Chung et al., 2008). El formulario final STOP-BANG consiste en ocho preguntas que incluyen las cuatro cuestiones originales del formulario STOP.

En el Anexo C se recoge un ejemplo de formulario STOP-BANG utilizado en clínica.

Ya que se ha demostrado la validez de este cuestionario para predecir SAHS (F Chung et al., 2012; F Chung, Liao, Sasaki, Elsaid, & Yegneswaran, 2011; Frances Chung et al., 2013; A Roebuck et al., 2014), y ha sido utilizado como complemento a la extracción de características de la señales de audio para mejorar la predicción de SAHS (Behar, Roebuck, Shahid, et al., 2013), se han incluido sus preguntas en los formularios entregados a los voluntarios.

Los formularios preparados constan de las siguientes preguntas:

- Fecha de nacimiento
- Altura
- Peso
- Perímetro del cuello
- Género
- ¿Es hipertenso?
- ¿Se le oye roncar con un volumen alto? (por ejemplo, se le oye a través de las puertas)
- ¿Se levanta cansado?
- ¿Le han observado que deje de respirar mientras duerme?
- ¿Le han diagnosticado que sufre de SAHS? ¿Está en tratamiento? ¿Cuál es el tratamiento?
- Dispositivo de grabación utilizado
- Ubicación del dispositivo de grabación
- Comentarios que puedan ser de interés.

3.1.3 Entrega de formularios y sistema de grabación al paciente

Una vez localizado el paciente, y acordada la fecha en la que realizará las grabaciones nocturnas, se le proporciona una grabadora MP3 modelo SONY ICD-UX200, así como pilas suficientes. Se le explica el manejo de la grabadora.

Si el paciente dispone de teléfono móvil con capacidad de grabar notas de voz, se le indica que también realice la grabación con su teléfono; y a ser posible, a la vez que con la grabadora, con el fin de obtener dos registros simultáneos con distintos dispositivos.

Se le proporcionan las indicaciones de cómo realizar las grabaciones (ubicación de los dispositivos) y los formularios que ha de rellenar.

Al final queda a discreción del voluntario con cuál dispositivo realiza la grabación.

3.1.4 Contestación a los formularios y realización de grabaciones nocturnas

En la bibliografía se recogen procedimientos de grabación del ronquido en el domicilio del paciente utilizando grabadoras MP3 (Kreivi, Salmi, Maasilta, & Bachour, 2013; Salmi & Virtanen, 2009) o teléfonos inteligentes (Behar, Roebuck, Shahid, et al., 2013).

Normalmente se ubican cerca de la cabeza del paciente (en la mesilla de noche o en la cabecera de la cama). Esta indicación de ubicación se les sugiere a los voluntarios.

Se registra el sonido a lo largo de toda la noche (7-8 horas). Normalmente, se pide al paciente que registre más de una noche.

La frecuencia de muestreo utilizada fue de 44KHz/Estéreo para la grabadora MP3. En el caso de los teléfonos móviles, la configuración es la fijada en la aplicación de notas de voz que viene por defecto con el aparato.

Un problema encontrado a la hora de ubicar los dispositivos de grabación fue la identificación de donde se encontraba el micrófono interno del dispositivo, para poder orientarlo hacia el paciente. La intensidad de la señal cambia bastante si el micrófono apunta hacia el paciente o en sentido contrario. A los voluntarios se les indicó que colocaran el micrófono del dispositivo apuntando hacia su cabeza, aunque finalmente queda a discreción del voluntario, al no poderse controlar este aspecto del registro.

En uno de los casos, se utilizó un micrófono externo omnidireccional de solapa, de calidad profesional, modelo RØDE smartLav. El tamaño del micrófono se puede observar en la Figura 3-4. En este caso, el micrófono se colocó en la cabecera de la cama, lo más cerca posible de la cabeza del paciente.

En algunos casos, el dispositivo de grabación se colgó de una lámpara sobre la cabeza del voluntario.

El voluntario rellena el formulario entregado en su casa. Se toma la medidas biométricas solicitadas él mismo, utilizando sus propios medios (cinta métrica, báscula de baño, etc.).

Con este entorno tan “poco controlado” se pretende emular las condiciones que se darían en el futuro caso de despliegue de una aplicación para *Smartphone* de amplia difusión.



Figura 3-4: Micrófono RØDE smartLav

3.1.5 Recogida de los formularios y grabaciones realizadas

Una vez que el paciente ha realizado las grabaciones, se recogen los dispositivos de grabación, los formularios respondidos y los ficheros con el audio.

Se realiza una primera revisión rápida de los ficheros de audio para verificar que son válidos (que se ha registrado algo y que la señal es adecuada) y que contienen ronquidos. En esta revisión rápida, se utilizó la herramienta de edición de audio “**Audacity®**”⁹. Es una herramienta gratuita y OpenSource disponible en múltiples plataformas.

3.1.6 Introducción en la BBDD de los datos del dispositivo de registro

Durante la investigación se han utilizado diferentes modelos de equipos para realizar las grabaciones. De nuevo, la elección de este modelo de trabajo se basa en emular los más posible el futuro despliegue de la aplicación para *Smartphone*. En esa futura investigación, cada paciente usará el *Smartphone* de su propiedad, con lo que el número de modelos distintos (y sus características de registro de audio) de estos dispositivos será alto. Se crea una entrada en la tabla ‘**grabadora**’ (ver anexo A) para cada tipo de dispositivo, recogiendo la mayor cantidad de características posibles. Los campos que se recogen son:

- Identificador único del tipo de equipo
- Marca
- Modelo
- Número de canales de registro
- Frecuencia de muestreo
- CODEC utilizado
- Compresión
- Formato
- Bit rate del stream (en Kbps)
- Frecuencia de corte en alta frecuencias

⁹ <http://sourceforge.net/projects/audacity/> y <http://web.audacityteam.org>

- Frecuencias de corte en baja frecuencias
- Sensibilidad del micrófono
- VOR
- LCF Low Cut
- Comentarios

En la Tabla 3-1 se recogen los sistemas de grabación utilizados.

MARCA	MODELO	CAN.	CODEC	B.R	FOR.	F.M(Hz)	COMENTARIOS
APPLE	iPhone 5S	1	AAC LC / CBR	64	M4A	44100	Usando el micrófono interno
SONY	ICD-UX200	2	MPEG Audio Layer 3 / CBR	192	MP3	44100	Grabadora MP3. Rango de frecuencia ST:10Hz-20.000Hz.
MOTOROLA	MOTO-G 16GB	1	MPEG Audio Layer 3 / CBR	32	MP3	44100	Usando el micrófono interno
APPLE	iPhone 5S + RØDE	1	AAC LC / CBR	64	M4A	44100	Micrófono externo RØDE smartLav
SAMSUNG	NOTE 4	1	AAC LC / CBR	128	M4A	44100	Usando el micrófono interno
LG	Optimus 3D P920	1	AAC LC / CBR	48	3GP	44100	Usando el micrófono interno

Tabla 3-1: Dispositivos de grabación utilizados y sus características

3.1.7 Introducción en la BBDD de los datos del paciente

Los datos recogidos en los formularios entregados a los voluntarios se introducen en la base de datos en la tabla *‘paciente’* (ver anexo A). Se crea una entrada para cada paciente. Los valores de algunos campos se calculan en base a los datos recogidos en los formularios. Los campos que se recogen son:

- Identificador único del paciente, asignado automáticamente por la BBDD
- “Nickname” del paciente
- Fecha de nacimiento
- Edad
- Género
- Altura
- Peso
- Índice BMI
- Perímetro del cuello
- Comentarios del paciente y del investigador

Según la OMS, el índice BMI se calcula como (World Health Organization (WHO), 2015):

$$BMI = \frac{\text{peso (kg)}}{(\text{altura (m)})^2} \quad (1)$$

También se informan campos booleanos recogidos de los formularios o derivados de los datos escalares:

- BMI ≥ 35
- Perímetro cuello ≥ 40
- Edad ≥ 50
- Ronca
- Se levanta cansado
- Es hipertenso
- Apneas observadas
- SAHS diagnosticado

3.1.8 Introducción en la BBDD de los datos de los registros realizados

Para cada fichero de audio recogido, se crea una entrada en la tabla ***‘audio’*** (ver anexo A) que recoge la descripción y características del registro realizado. Los campos que se recogen son:

- Identificador único del registro de audio, asignado automáticamente por la BBDD
- Identificador del paciente
- Fecha de registro
- Hora de inicio
- Duración
- Duración en segundos
- Identificador de la grabadora
- Número de canales
- Frecuencia de muestreo
- Compresión utilizada
- Bit rate
- Número de muestras
- Comentarios
- *OID del Fichero de audio*
- *Nombre del Fichero de audio*

Los campos “*OID del fichero de audio*” y “*Nombre del fichero de audio*” no se informan en este paso. Los valores correspondientes se informan al insertar el fichero de audio original en la BBDD, tal y como se describe en el apartado 3.1.9.

3.1.9 Introducción en la BBDD de los ficheros de audio originales

Una vez creado el registro de audio, se carga en la BBDD el fichero original. Para ello se utiliza el mecanismo de *LARGE OBJECTS* de PostgreSQL. Este mecanismo permite almacenar registros con un tamaño máximo de 4TB.

Al insertar el fichero en la tabla de Large Objects, la BBDD devuelve un OID que se almacena en el campo “*OID del fichero de audio*” de la tabla ***‘audio’*** (ver anexo A) del registro creado en el paso anterior (3.1.8.). Así mismo, también se incluye en el campo “*Nombre del fichero de audio*” de la tabla ***‘audio’*** el nombre del fichero subido.

Los ficheros de audio originales se archivan en disco como se describe en la sección 3.4.1.

Hay definida una restricción sobre estos campos, de manera que tanto el OID como el nombre del fichero han de ser únicos.

3.2 Normalización y extracción de características

El siguiente paso a la obtención de los registros de audio es su normalización para su análisis.

Se seleccionan segmentos significativos de los ficheros de audio normalizados, y se procede a la extracción de características de estos segmentos que sirvan para clasificar los eventos respiratorios. Así mismo, se realiza una clasificación manual de los eventos respiratorios de los segmentos. Esta clasificación manual de los eventos proporciona los conjuntos de datos de entrenamiento y test para los modelos matemáticos de clasificación que se utilicen.

Todas las características calculadas y clasificaciones manuales realizadas son introducidas en la base de datos para su posterior tratamiento y análisis.

Este subproceso queda reflejado en la Figura 3-5.

A continuación se detallan las tareas que componen este subproceso.

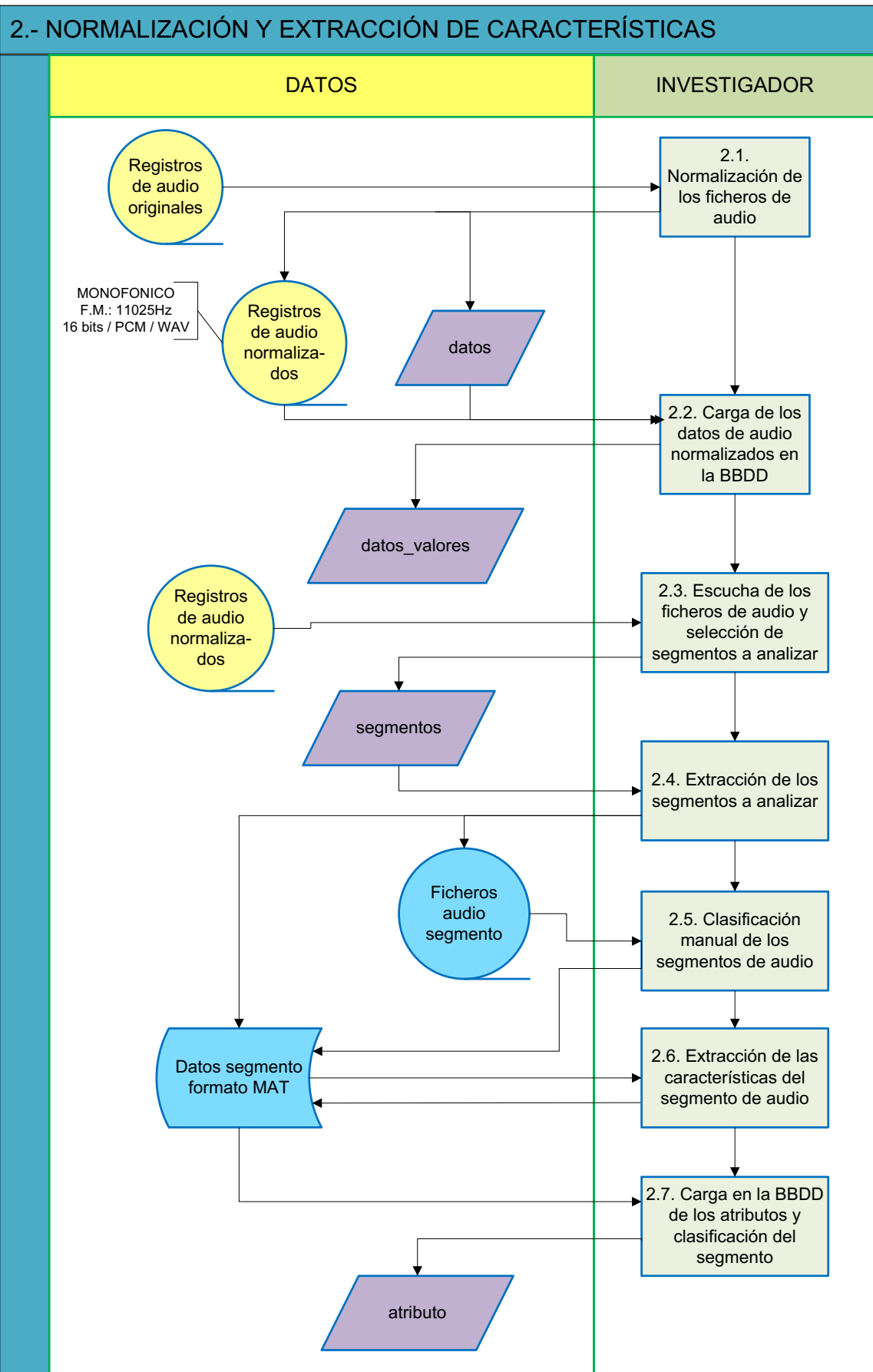


Figura 3-5: Descripción del subproceso de normalización y extracción de características

3.2.1 Normalización de los ficheros de audio

Debido que durante el proceso de adquisición de los registros de audio se han utilizado diferentes dispositivos de grabación, los que implica obtener ficheros con diferentes frecuencias de muestreo, número de canales y sistema de codificación, se realiza una normalización de los registros de audio originales de la siguiente manera:

1. Utilizando Audacity®, se realiza la mezcla de canales de registros estereofónicos a un canal monoaural. Se les da la misma ganancia a ambos canales (media de los dos canales). En el caso de registros monoaurales no se realiza este paso.
2. Inicialmente, en este trabajo, no se aplica ningún tipo de filtro de reducción de ruido.
3. Se submuestra la señal a 11025Hz (Behar, Roebuck, Shahid, et al., 2013; Dafna, Tarasiuk, & Zigel, 2013; Snider & Kain, 2013), ya que las frecuencias que nos interesan están por debajo de 4 KHz (Fiz Fernández, Solà Soler, & Jané Campos, 2011; Pevernagie, Aarts, & De Meyer, 2010; Yadollahi & Moussavi, 2010). La frecuencia fundamental del ronquido está por debajo de los 1000 Hz (Fiz et al., 1996).

Audacity® utiliza la librería **libresample**¹⁰ para el remuestreo de las señales de audio. Esta librería es la implementación del algoritmo de remuestreo usando la función “*sinc*”, escrita originalmente por Dominic Mazzoni basándose en el código de Julius Smith de la Universidad de Stanford (Smith, 2002).

4. La señal resultante se graba en formato de audio WAV/PCM/16 bits.

A la señal normalizada obtenida en este paso, la denominamos *s* para su posterior referencia.

3.2.2 Carga de los datos de audio normalizados en la BBDD

Los metadatos de *s* se incorporan a la BBDD a través de la tabla ‘*datos*’ (ver anexo A). Los campos que se recogen son:

- Identificado del registro de audio original
- Identificador único del fichero de audio normalizado
- Número de canales
- Resolución en bits
- Nombre del fichero en disco (nombre único) (ver sección 3.4.1)
- Duración en segundos
- Filtro aplicado
- Frecuencia de corte superior
- Frecuencia de corte inferior
- Comentarios de utilidad

¹⁰ <http://ftp.debian.org/pool/main/libr/libresample>

El fichero obtenido WAV obtenido en el paso 3.2.1 se carga en la base de datos en la tabla ‘*datos_valores*’ (ver anexo 0). Los campos que se recogen son:

- Identificador único del fichero de audio normalizado
- Posición en la secuencia temporal del valor a introducir
- Valor de la posición temporal

Para realizar la carga masiva de datos de audio se probaron distintos método y herramientas. La adoptada finalmente (en base a un menor tiempo de procesado) consiste en generar un fichero CSV a partir de los datos PCM del fichero WAV, y cargarlo en la BBDD con la función “*Import*” de la utilidad de administración de PostgreSLQ “*pgAdmin-III*”.

Se ha desarrollado en Matlab® la función **wav2csv** para convertir el fichero WAV al formato CSV con la estructura adecuada para su carga con “*pgAdmin-III*”. Esta función se recoge en el anexo A.

Los datos de valores de la posición temporal (secuencia de audio) se almacenan en formato de enteros de 16 bits (formato nativo PCM/WAV) en lugar de almacenarlos en un formato de coma flotante. Esto permite reducir el espacio necesario para almacenamiento en cuatro veces (el formato *double* ocupa 64 bits).

Así mismo, el fichero WAV se archiva en disco según se describe en la sección 3.4.1.

3.2.3 Escucha de los ficheros de audio y selección de los segmentos a analizar

La señal de audio normalizada *s* se revisa manualmente, escuchando sus contenidos. De nuevo, la herramienta utilizada es Audacity®. Se seleccionan segmentos que recojan los eventos que se están buscando y que no contengan mucho ruido ambiental (sonidos del paciente al acostarse, actividades nocturnas en la casa, ruido de vehículos del exterior, etc.).

Aunque en un futuro la perturbación por ruidos no respiratorios es un aspecto que habrá que tener en cuenta, para la realización de este trabajo se han descartado estas secciones “ruidosas”, seleccionando solo segmentos con una identificación clara de los eventos buscados.

Como ayuda se visualiza simultáneamente la forma de onda y el espectrograma de la señal (Figura 3-6).

El espectrograma es el resultado de calcular el espectro de tramas enventanadas de una señal. Resulta una gráfica tridimensional que representa la energía del contenido de frecuencias de la señal según va variando ésta a lo largo del tiempo. Para calcular el espectro se utiliza la Transformada Rápida de Fourier (FFT).

Para la visualización del espectrograma de señal, se configuran los siguientes parámetros en Audacity®:

- Ventana FFT:
 - Tamaño de la ventana: 2048
 - Tipo de ventana: Hamming
- Mostrar:
 - Frecuencia mínima: 0 Hz
 - Frecuencia máxima: 4 KHz
 - Ganancia: 20 dB
 - Rango: 120 dB
 - Ganancia de frecuencia: 0 dB/dec

En la Figura 3-6 se puede observar la visualización simultánea de señal y espectrograma. Ambas vistas están sincronizadas en el tiempo (eje X de la gráfica).

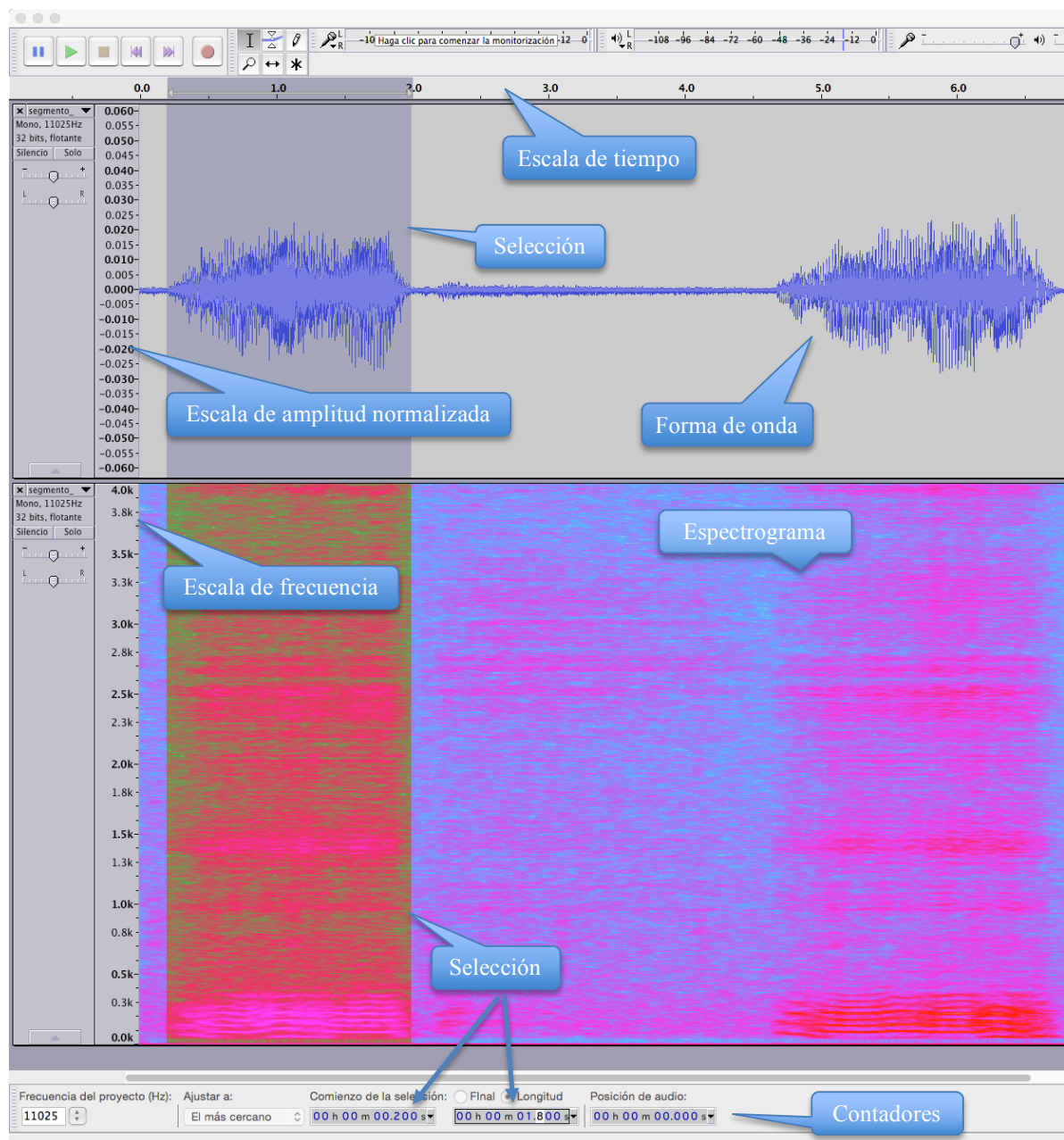


Figura 3-6: Captura de pantalla de la aplicación Audacity® mostrando la forma de onda y su espectro de un registro de ronquidos.

La gráfica superior se corresponde con la forma de onda de la señal, normalizada a $[-1 \ 1]$. Su escala de la izquierda refleja los valores de amplitud de la señal.

La gráfica inferior corresponde al espectrograma de la señal calculado y visualizado con los parámetros indicados anteriormente. La escala vertical indica la frecuencia. Se selecciona un tamaño de ventana de 2048 muestras para aumentar la resolución en frecuencia.

La intensidad del color (tercer eje de la gráfica) indica la concentración de energía en esa frecuencia. Se controla la visualización con los valores de ganancia y rango. El espectrograma muestra seis bandas de colores de mayor a menor energía: blanco – rojo – magenta – azul oscuro – azul claro – gris. Con la configuración que se ha definido, todos los valores por encima de -20dB (máxima energía) aparecerán en blanco, mientras que los valores por debajo de -120dB (mínima energía) aparecerán en gris.

Se puede hacer zoom tanto de la escala horizontal del tiempo (ambas gráficas simultáneamente), como de las verticales (cada gráfica de forma independiente). Los límites del segmento se definen de forma arbitraria en base a segundos enteros, tanto en su inicio como en su duración.

Usando los contadores de tiempo de la parte inferior de la ventana de visualización de la aplicación (Figura 3-6), es muy sencillo obtener los valores de inicio, final y duración del segmento seleccionado.

A los segmentos obtenidos en este paso, les denominamos $s(n)$, siendo n el índice del segmento extraído.

Los metadatos de $s(n)$ se introducen en la tabla '**segmento**' (ver anexo 0). Los campos que se recogen son:

- Identificador del fichero de audio normalizado del que procede
- Identificador único del segmento $s(n)$
- Tiempo de inicio
- Segundo de inicio
- Duración en segundos
- Segundo de finalización
- Indicador de si contiene eventos SAHS (apneas)
- Ventana a aplicar en la clasificación y extracción de características (ver sección 3.2.5)
- Número de muestras en el segmento
- Número de patrones a obtener del segmento
- Tipo de filtro aplicado
- Frecuencia superior de corte
- Frecuencia inferior de corte
- Comentarios de utilidad

En este trabajo, no se ha aplicado ningún tipo de filtrado a los segmentos $s(n)$ seleccionados.

3.2.4 Extracción de los segmentos a analizar

Una vez definidos los segmentos $s(n)$ se extraen sus ficheros de audio y datos directamente de la BBDD utilizando una función creada en Matlab®. Esta función (**extraer_segmento**, recogida en el anexo A) toma como argumento el identificador del segmento a extraer y genera dos ficheros de salida:

- Un fichero WAV con el audio del segmento especificado
- Un fichero de Matlab® binario en formato MAT, con los datos de audio y metadatos del segmento, así como todas las matrices necesarias para los pasos de clasificación (sección 3.2.5) y extracción de características (sección 3.2.6).

Los nombres de los ficheros los genera la función a partir del identificador de fichero, para que no haya error a la hora de identificarlos.

En la BBDD los datos de audio de los ficheros WAV normalizados (secciones 3.2.1 y 3.2.2) están almacenados en formato nativo (enteros de 16 bits). Para su análisis en Matlab®, se convierten a valores en coma flotante en el rango $[-1 \ 1]$, ya que esto permite comparar fuentes sonoras que tengan distinto formato numérico en su codificación (enteros de 8, 16, 24 o 32 bits, coma flotante, etc.). Esta conversión es realizada por la función **extraer_segmento** al recuperar los datos de la BBDD.

También se ha generado una función para la descarga el lote de todos los segmentos generados a partir de un fichero de datos de audio (función **extraer_segmento_todos**, recogida en el anexo A).

3.2.5 Clasificación manual del segmento de audio

Cada segmento $s(n)$ se divide en bloques (enventanado) utilizando una ventana rectangular $w(n)$ de longitud N muestras. De esta manera, nombramos la k -ésima ventana obtenida como $s_k(n)$.

La clasificación manual de eventos de $s(n)$ se realizó en base a la percepción acústica y visual de evento y su contexto. Se escucha y visualiza el segmento $s(n)$ y se asigna, a criterio de la persona que realiza la clasificación, una de las tres clases definidas (eventos respiratorios) a cada bloque $s_k(n)$:

- Silencio,
- Respiración
- Ronquido.

Para la clasificación se utilizó como ayuda la herramienta de audio Audacity®. La percepción acústica se logra con la reproducción del segmento, y como ayuda de la percepción visual se visualiza el espectrograma y la forma de onda simultáneamente, tal y como se puede observar en la Figura 3-6. Con la ayuda de los contadores de tiempo, es relativamente sencillo llevar el control de la ventana $s_k(n)$. Las asignaciones de clase para cada bloque $s_k(n)$ se apuntan (esto conforma el vector de clasificación $Clase_k(n)$) para

posteriormente añadirse a la base de datos. Todas estas operaciones se realizan manualmente.

Para cada $s_k(n)$ asignamos su clase, que denominamos $Clase_k$.

En este trabajo, la función ventana de análisis $w(n)$ tiene una duración de 100 ms. Teniendo $s(n)$ una frecuencia de muestreo de 11025 Hz, se ha utilizado $N=1102$ muestras.

En la bibliografía encontramos artículos que aplican ventanas $w(n)$ desde los 20ms (Yadollahi & Moussavi, 2010) hasta 200ms (Dafna et al., 2013). En el presente trabajo, el tamaño de $w(n)$ elegido viene determinado por las siguientes restricciones:

- Si $w(n)$ es muy pequeña, el número de las operaciones manuales de clasificación aumentan, por lo que es más fácil cometer un error.
- Debido a como se realiza la clasificación manual con la herramienta Audacity®, es de bastante ayuda que el tamaño de la ventana sea en potencias de 10 (1ms – 10ms – 100ms – 1s) para el manejo de los contadores.
- Si $w(n)$ es muy grande, aumenta la probabilidad de tener dos tipos de evento en la misma ventana, por lo que aumentamos el ruido de la clasificación al introducir una incertidumbre adicional al técnico que está realizando la clasificación.

En la práctica diaria de este trabajo, se ha comprobado que elegir $w(n) = 100ms$ es un compromiso aceptable para las restricciones expuestas.

Un problema encontrado la hora de realizar la clasificación fue establecer un criterio claro y objetivo para definir los eventos respiratorios en las fronteras entre ellos. Sobre todo en diferenciar cuando se produce una respiración fuerte frente a un ronquido débil, o una respiración débil frente al ruido de fondo. No se ha encontrado en la bibliografía un método objetivo para guiar al operador en esta separación, por lo que se dejó a criterio del técnico que clasificaba los segmentos de audio.

El vector de clasificación $Clase_k(n)$ obtenido, se incorpora manualmente al fichero de datos de $s(n)$ en formato Matlab® obtenido en el paso 3.2.4.

3.2.6 Extracción de las características del segmento de audio

Podemos asumir que los sonidos respiratorios durante el sueño se producen de la misma manera que los fonemas de la voz (vibración de las estructuras de la vía aérea superior), por lo que podemos aplicar técnicas de reconocimiento de voz (A Roebuck et al., 2014). Así mismo, al analizar los sonidos respiratorios, podemos abordarlo desde el punto de vista espectral o desde el dominio del tiempo (Dafna et al., 2013).

Tras revisar los distintos métodos publicados, se observa que los mejores resultados se obtienen usando características de ambos dominios (Karunajeewa et al., 2008), por lo que se seleccionan cuatro características del audio para su clasificación que mejor se adapten a las características de este trabajo (bajo coste computacional y posibilidad de calcularse on-line):

1. Zero Crossing Rate (Z_k)
2. Short Time Energy (Er_k)

3. Coeficiente de correlación normalizado a 1ms de retardo (C_k)
4. Primer coeficiente predictor de análisis LPC de 12 polos (Lp_k)

Para el cálculo de estos atributos de audio, se ha desarrollado una función en Matlab® llamada **analizar_segmento**, según se recoge en el anexo A. Los datos para este análisis se extraen del fichero Matlab® generado en el paso 3.2.4.

A continuación se analizan las propiedades y posibilidades de cada una de las características seleccionadas.

3.2.6.1 Zero Crossing Rate (Z_k)

En el contexto de señales discretas en el tiempo, se dice que se produce un cruce por cero cuando muestras sucesivas tienen diferente signo algebraico (Shete & S.B. Patil, 2014). El ratio de cruce por cero (*Zero Crossing Rate*) es una medida del número de veces que en un intervalo de tiempo/ventana se produce un paso por cero de la amplitud de la señal. Esta medida nos da una estimación inicial de las propiedades espectrales de la señal con un coste computacional muy bajo.

Esta característica ya se utilizó en 1962 por Sakai y Doshita, de la Universidad de Kyoto, para construir un sistema de reconocimiento de vocales/consonantes (Benesty, Sondhi, & Huang, 2008).

Zero Crossing Rate es un parámetro importante para la clasificación de voz/no-voz, ya que es un indicador de la frecuencia a la cual la energía se concentra en el espectro de la señal (Bachu, Kopparthi, Adapa, & Barkana, 2010).

La definición de Zero Crossing Rate Z_k es:

$$Z_k = \frac{1}{N} \sum_{m=1}^N |sgn[s_k(m)] - sgn[s_k(m-1)]| \quad (2)$$

donde

$$sgn[s_k(n)] = \begin{cases} 1, & s_k(n) \geq 0 \\ -1, & s_k(n) < 0 \end{cases} \quad (3)$$

Aplicando esta medida a cada ventana $s_k(n)$, obtenemos Z_k que representa el *Zero Crossing Rate* de la ventana.

En la Figura 3-7 podemos observar la representación de Z_k obtenido de un pequeño segmento de audio. En la figura también se representa la forma de onda (amplitud en el tiempo) y la clase asignada a cada ventana $s_k(n)$.

Partiendo del hecho que el ronquido presenta una estructura similar a la voz, una generalización razonable es que si Z_k es alto, la señal no presenta estructura, mientras que si es bajo, presenta una estructura similar a la locución (Bachu et al., 2010; Karunajeewa et al., 2008), por lo que nos podría servir para separar ronquidos del resto de eventos.

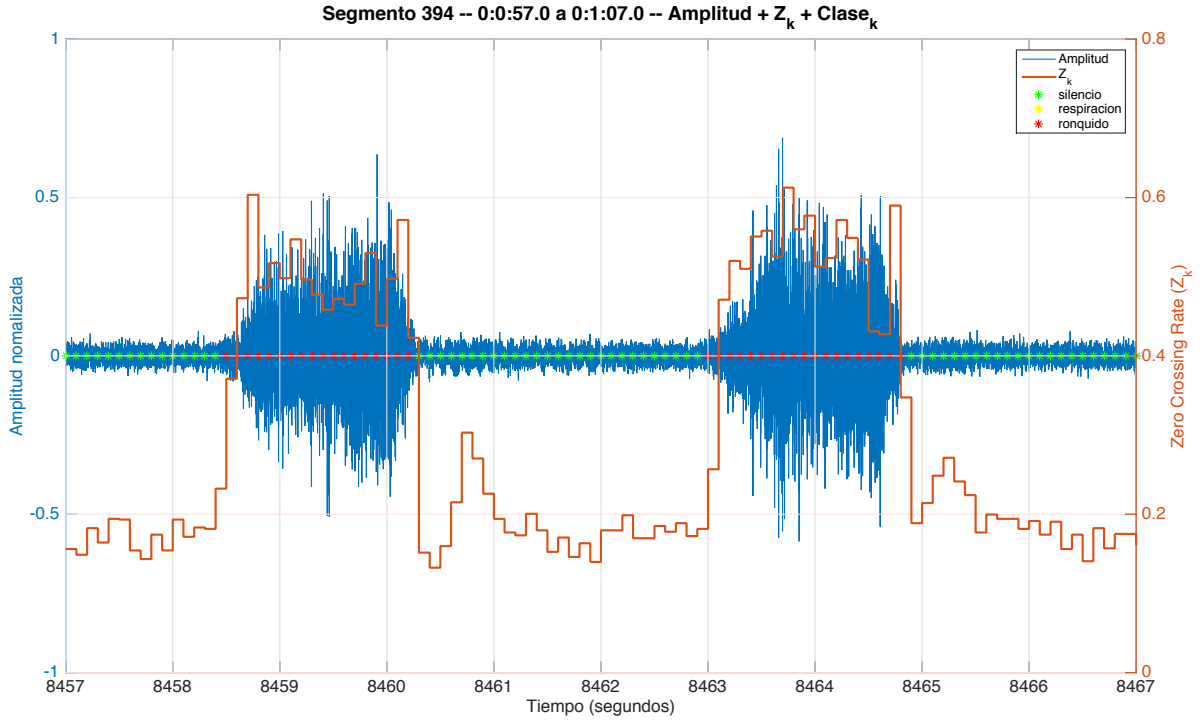


Figura 3-7: Representación del Z_k calculado, forma de onda y clase asignada para una pequeña porción de un segmento de audio

3.2.6.2 Short Time Energy (Es_k)

La energía de la voz es otro parámetro para la clasificación de voz/no-voz. La sección con voz de la señal tiene una alta energía debido a su periodicidad, mientras que las secciones de no-voz tienen baja energía (Bachu et al., 2010).

La energía de la señal en decibelios de cada ventana $s_k(n)$, se calcula como:

$$Es_k = 10 \log \left(\varepsilon + \frac{1}{N} \sum_{m=1}^N (s_k)^2(m) \right) \quad (4)$$

donde ε es una constante positiva pequeña añadida arbitrariamente para prevenir el error de calcular el logaritmo de cero (Karunajeewa et al., 2008).

En la Figura 3-8 podemos observar la representación de Es_k obtenido de un pequeño segmento de audio. En la figura también se representa la forma de onda (amplitud en el tiempo) y la clase asignada a cada ventana $s_k(n)$.

Esta es una característica que depende de los parámetros de grabación, como la sensibilidad del micrófono o la distancia a la fuente de sonido, por lo que es una medida insuficiente por ella misma para detectar ronquidos. Sin embargo, presenta una gran capacidad de discriminación entre eventos de voz/no-voz, siendo los ronquidos los eventos con mayor energía y los silencios los que presentarían unos niveles más bajos (Karunajeewa et al., 2008).

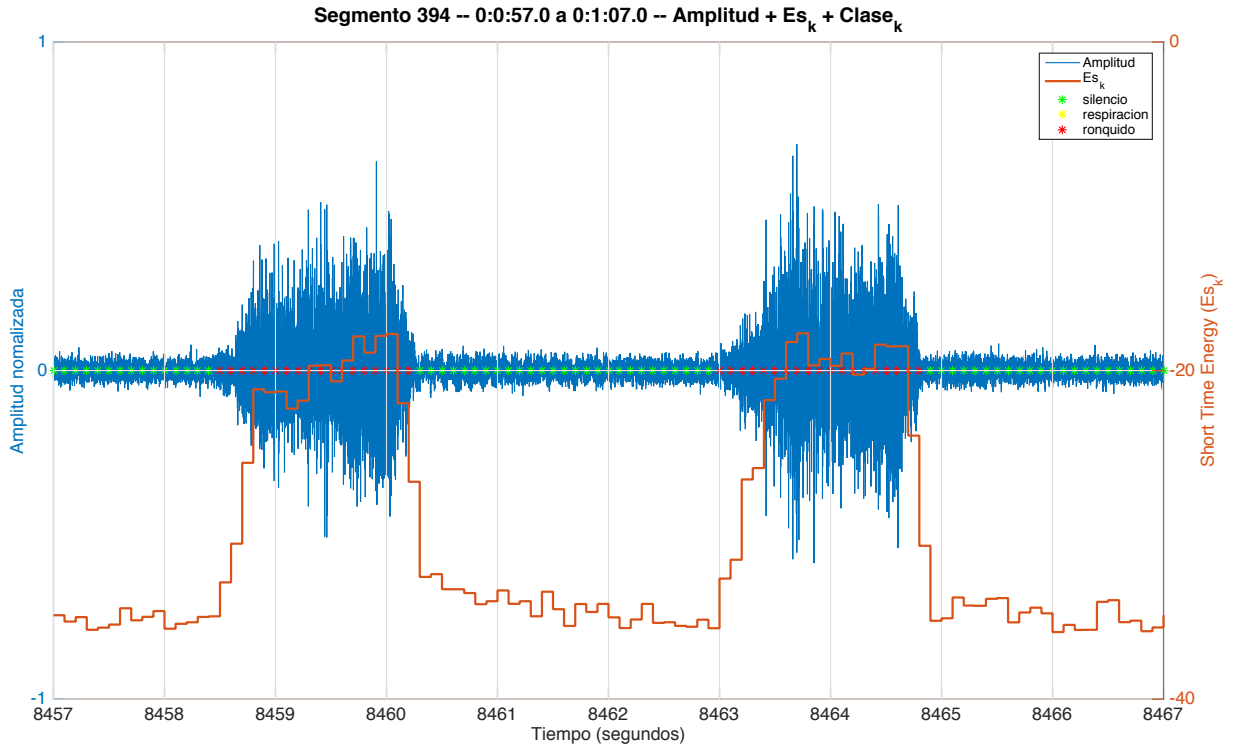


Figura 3-8: Representación de Es_k calculado, forma de onda y clase asignada para una pequeña porción de un segmento de audio

3.2.6.3 Coeficiente de autocorrelación normalizado a 1ms de retardo (C_k)

Las secciones de señal que contengan estructura (ronquidos/voz) deberían presentar una correlación más alta entre muestras consecutivas que las secciones sin señal, debido a la concentración de energía de baja frecuencia (Karunajeewa et al., 2008).

C_k se calcula como

$$C_k = \frac{\sum_{n=1}^N s_k(n)s_k(n - \frac{fs}{1000})}{\sqrt{(\sum_{n=1}^N s_k^2(n))(\sum_{n=0}^{N-1} s_k^2(n))}} \quad (5)$$

En la Figura 3-9 y Figura 3-10 podemos observar la representación de C_k obtenido de un pequeño segmento de audio. En las figuras también se representa la forma de onda (amplitud en el tiempo) y la clase asignada a cada ventana $s_k(n)$.

La elección de un retardo de 1ms se extrajo de los resultados experimentales de Karunajeewa et al.(2008), en los que este desplazamiento proporcionaba la máxima discriminación entre las clases. En futuros trabajos habrá que analizar si esta retardo es el más óptimo para un entorno heterogéneo de sistemas y entornos de registro.

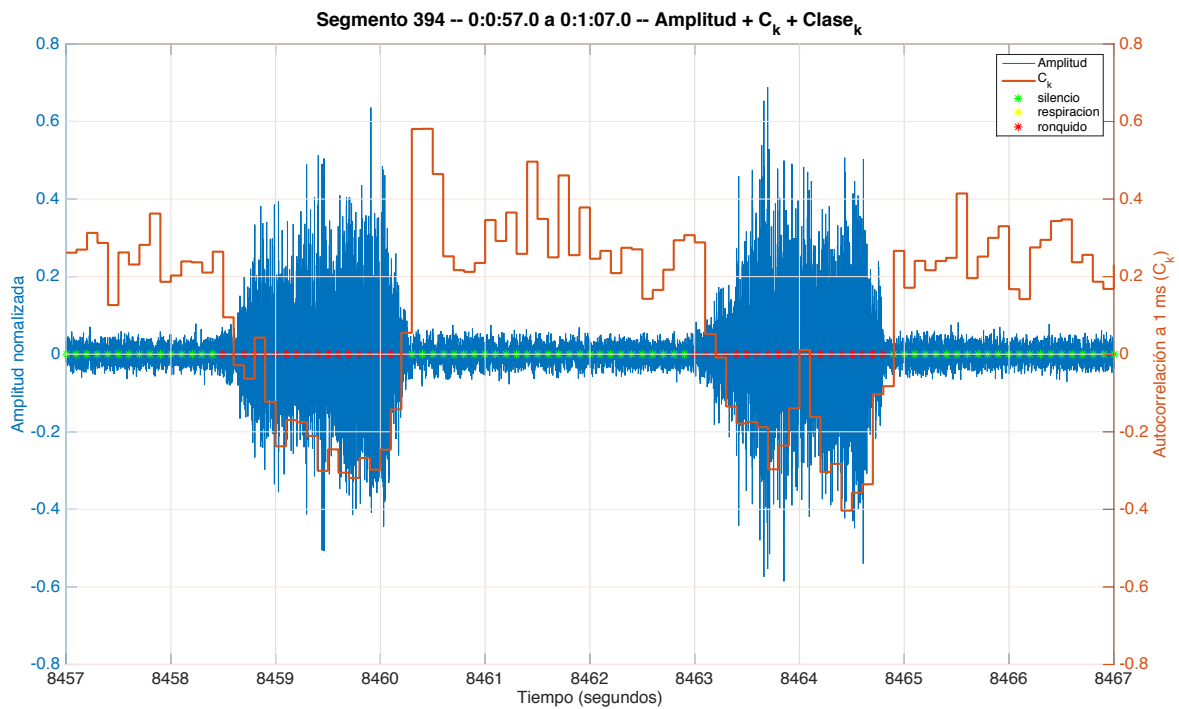


Figura 3-9: Representación de C_k calculado, forma de onda y clase asignada para una pequeña porción de un segmento de audio

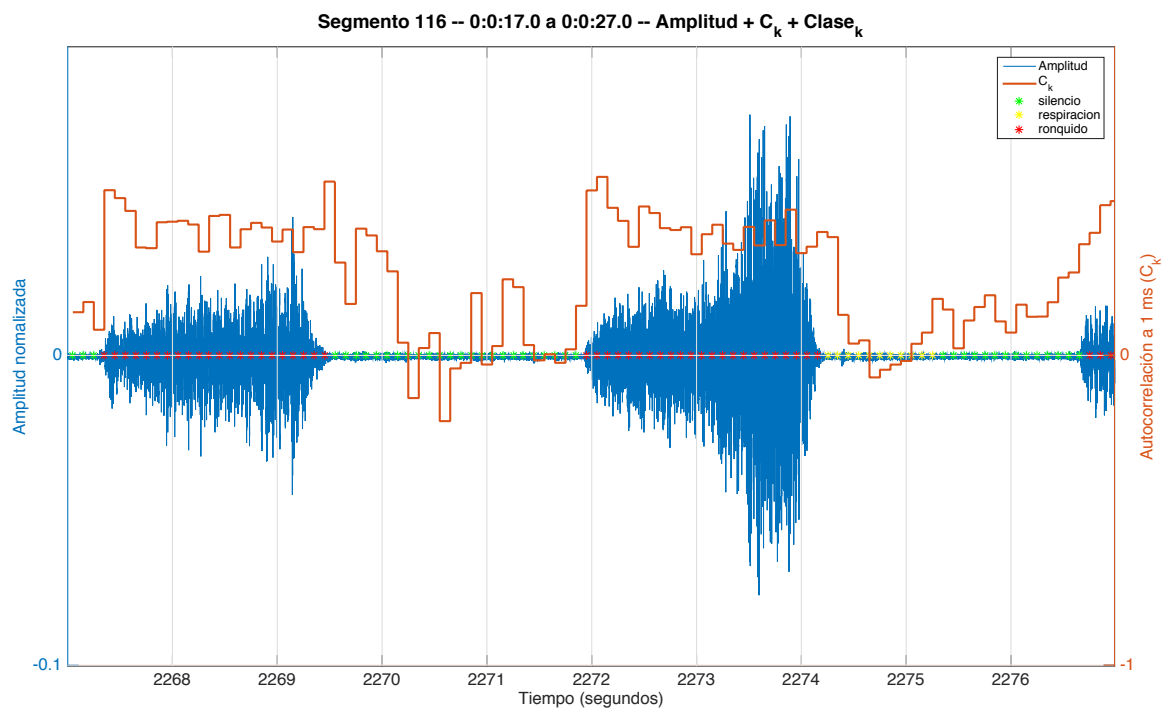


Figura 3-10: Representación de C_k calculado, forma de onda y clase asignada en otro segmento de audio

Esta medida está muy influenciada por los sonidos de fondo del ambiente, sobre todo si contienen estructura, tal y como se comprueba al comparar ambas figuras (Figura 3-9 y Figura 3-10).

3.2.6.4 Primer coeficiente predictor de análisis LPC de 12 polos (Lp_k)

Linear Predictive Coding (LPC) es un método de modelado de señales utilizado en procesamiento de voz.

Parte de la premisa de que existe fuerte correlación entre muestras consecutivas, especialmente en los tramos sonoros, que presentan una alta periodicidad. En base a esta premisa, podemos expresar la muestra actual aproximada como una combinación lineal de muestras anteriores:

$$\tilde{s}_k(n) = \sum_{k=1}^p \alpha_k \cdot s(n - k) \quad (6)$$

donde $\{\alpha_k\}$ son los coeficientes de predicción lineal, $s(n)$ es la señal real y $\tilde{s}_k(n)$ es la señal predicha.

LPC se utiliza como herramienta para la extracción de formantes. Es más adecuado para el modelado de vocales ya que son periódicas, excepto las vocales nasales. El método se basa en la función filtro (resonadores) del tracto vocal (Snell & Milinazzo, 1993). En Figura 3-11 la se representa un modelo de tubos resonantes del tracto vocal.

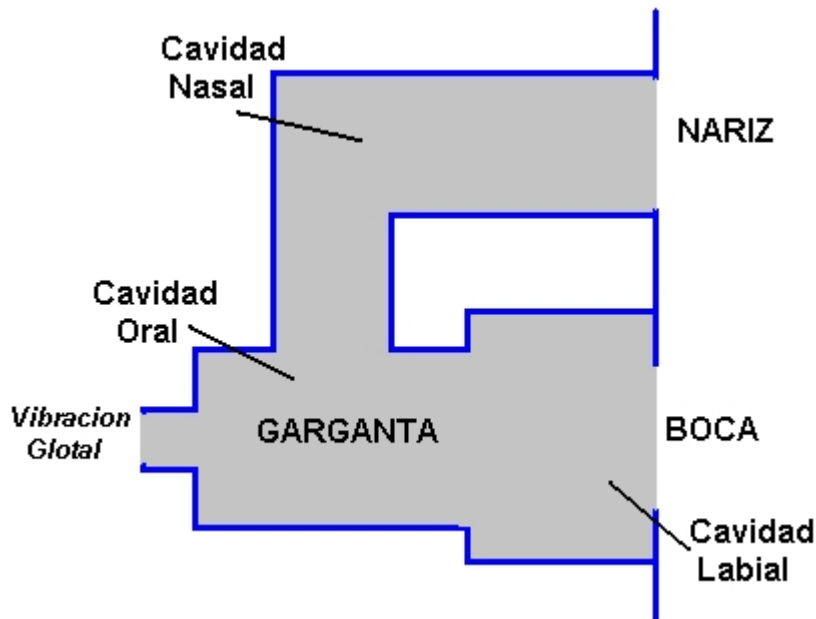


Figura 3-11: Modelo de tubos resonantes del tracto vocal

Cada tubo resonante funciona como un filtro acústico en función de su diámetro y longitud. Estos dos parámetros definen cuál es la frecuencia fundamental (o de resonancia) del tubo. Es lo que también se denomina formante. Los filtros se caracterizan por sus polos y ceros, es decir, cómo se comporta frente a la entrada de una señal sinusoidal, representando la respuesta sobre una circunferencia unidad. En los polos de un filtro es donde se produce la mayor amplificación de la señal de entrada, en oposición a sus ceros, que es donde se produciría la mayor atenuación. Para que los coeficientes del sistema sean reales, los ceros y polos aparecerán como pares complejos conjugados.

En general, la concatenación de N tubos resonadores da como resultado un sistema de N -polos. Para una concatenación de N tubos, hay al menos $N/2$ polos conjugados complejos, o resonancias o formantes. Experimentalmente se ha observado que el tracto vocal tiene aproximadamente un formante por kilohercio (Huan, Acero, & Hon, 2001).

Para especificar el orden del modelo, se ha seguido la regla general de dos veces el número de formantes esperados más dos. Debido a la frecuencia de muestreo elegida (ver sección 3.2.1) de 11 KHz, el corte de alta frecuencia lo hemos fijado en 5 KHz, por lo que es de esperar que tengamos 5 formantes en la señal, a razón de un formante por KHz. Por lo tanto, realizaremos el análisis LPC con 12 polos ($5*2+2$).

En la Figura 3-12 podemos observar la representación de Lp_k obtenido de un pequeño segmento de audio. En la figura también se representa la forma de onda (amplitud en el tiempo) y la clase asignada a cada ventana $s_k(n)$.

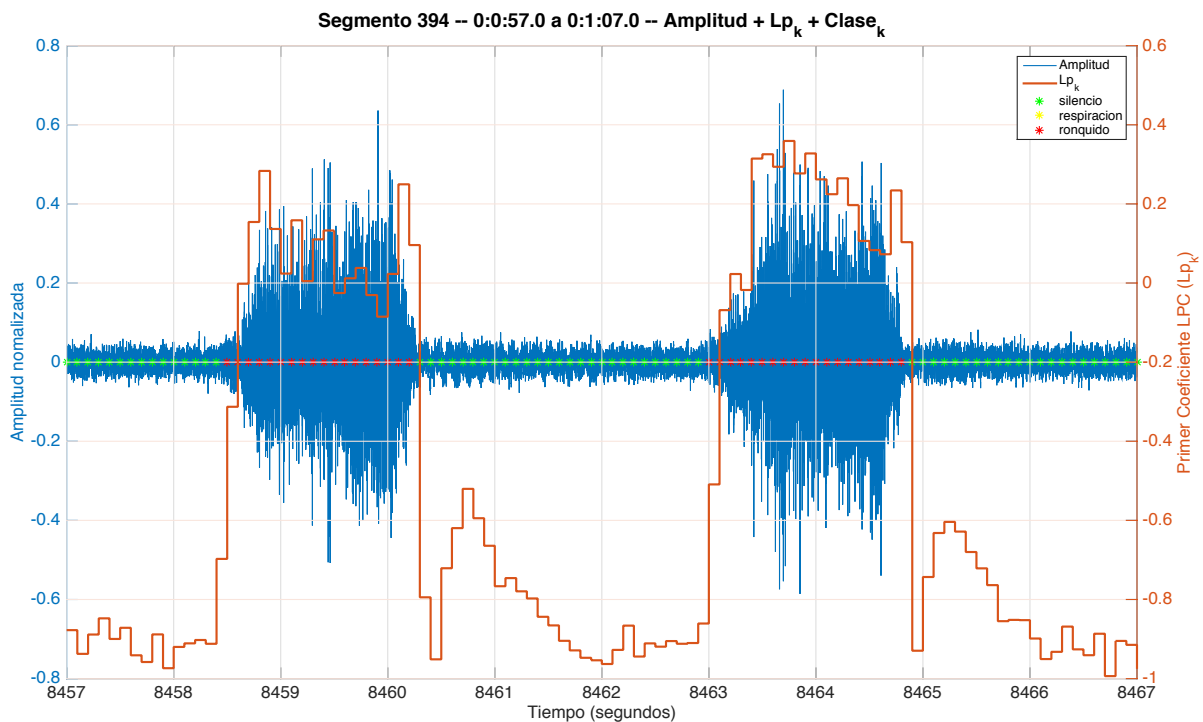


Figura 3-12: Representación de Lp_k calculado, forma de onda y clase asignada para una pequeña porción de un segmento de audio

Oppenheim & Schafer (1968) encontraron que el primer coeficiente LPC es idéntico al negativo del cepstrum con un retardo de una unidad de muestreo. Ya que el espectro de las tres clases que buscamos es diferente, es de esperar que el primer coeficiente LPC también sea considerablemente diferente para cada clase (Karunajeewa et al., 2008).

3.2.7 Carga en la BBDD de los atributos y clasificación del segmento

Para cada bloque $s_k(n)$ se genera un vector de características y clase:

$$f_k = \{Z_k \ Es_k \ C_k \ Lp_k \ Clase_k\}^t \quad (7)$$

El cálculo de los atributos y la generación del vector f_k se realiza a través de una función creada en Matlab®, denominada **analizar_segmento** (recogida en el anexo A).

El conjunto de vectores calculados para $s(n)$ se inserta en la base de datos en la tabla '**atributo**' (ver anexo 0). Los campos que se recogen son:

- Identificador del segmento $s(n)$
- Tiempo (t)
- Z_k
- C_k
- Es_k
- Lp_k
- $Clase_k$

Para la carga de todos estos datos, se ha desarrollado una función de Matlab®, denominada **subir_atributos** (recogida en el anexo A). Esta función lee los datos almacenados en el fichero MAT y los carga en la tabla '**atributo**'.

3.3 Selección del modelo de clasificación

Una vez obtenidos y tabulados todos los datos, se procede a su análisis y elección del clasificador que mejor resultados de en cuanto a las necesidades del proyecto.

Las herramientas elegidas para realizar el análisis de los datos son WEKA¹¹ versión 3.6.12 (Hall et al., 2009) para la clasificación de los patrones obtenidos, y Matlab® versión R2015a para el análisis y manejo de los datos.

Este subproceso queda reflejado en la Figura 3-13.

¹¹ <http://www.cs.waikato.ac.nz/~ml/weka/>

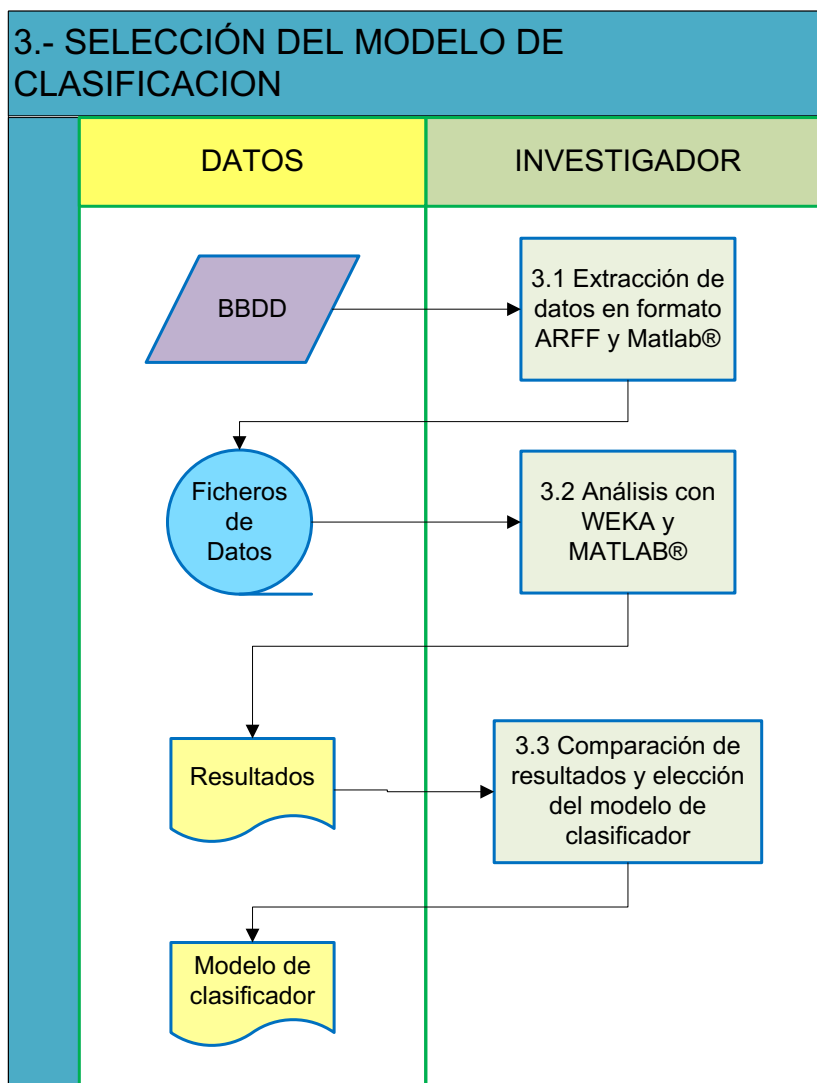


Figura 3-13: Descripción del subproceso de selección del modelo de clasificación

A continuación se detallan las tareas que componen este subproceso.

3.3.1 Extracción de datos en formato ARFF y MATLAB®

Antes de realizar el análisis, es necesario extraer los datos de la BBDD. La extracción se realiza directamente desde WEKA (Hall et al., 2009) y Matlab® a través de una conexión JDBC, o desde pgAdmin mediante consultas SQL que generan ficheros CSV.

En Matlab® se utiliza la aplicación *Database Explorer* para extraer los datos y generar las matrices necesarias para su análisis. En la Figura 3-14 podemos observar una captura de pantalla de esta herramienta.

Una vez configurada la conexión JDBC a la BBDD, con *Database Explorer* se pueden seleccionar las tablas a utilizar, los campos a extraer y especificar filtros (sentencias WHERE) y uniones de tablas (sentencias JOIN) mediante un interfaz gráfico. Los datos obtenidos se pueden almacenar en estructuras tipo ARRAY, CELL ARRAY, STRUCTURE, TABLE o DATASET, según las necesidades del proyecto.

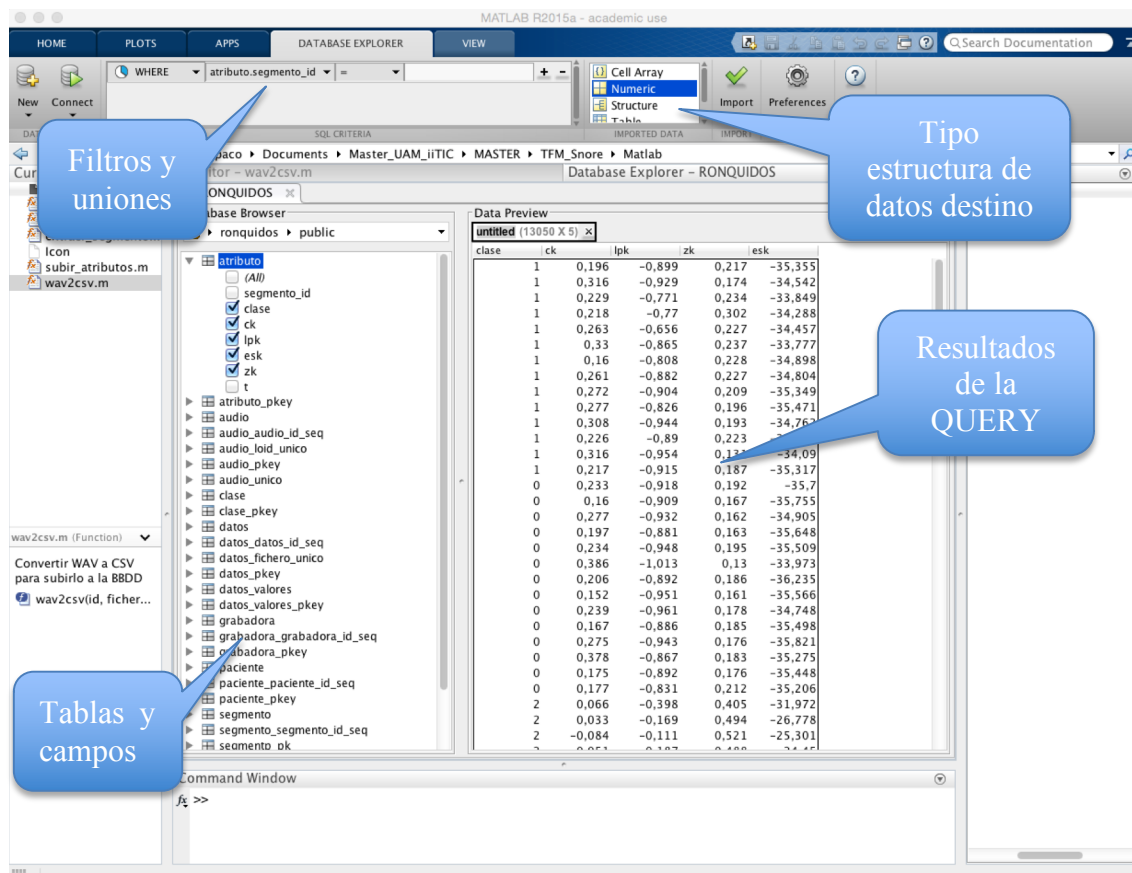


Figura 3-14: Matlab® Database Explorer

Se ha desarrollado una función en Matlab® llamada **extraer_atributos_segmento** que extrae en ARRAYS los valores de atributos, clase y características del segmento $s_k(n)$ indicado.

En WEKA, usando el GUI, desde el Explorer podemos conectarnos a bases de datos y realizar extracciones de datos. En la Figura 3-15 podemos ver como es el interfaz de conexión a base de datos.

Se observa que el interfaz de WEKA es menos elaborado que el de Matlab®, y que requiere de un mayor conocimiento del esquema de la BBDD a utilizar, que que hay que crear las consultas de selección a mano, sin un asistente.

Una vez descargados los datos desde la BBDD, se pueden almacenar localmente en un fichero ARFF.

Para facilitar la extracción de los datos, se ha generado la vista “*atributo_full*” (ver anexo A) en PostgreSQL, que realiza la unión de todas las tablas implicadas. Se ha realizado la reasignación de tipos en algunos campos para que el extractor de WEKA los considere como nominales. WEKA considera que son nominales todos los campos no numéricos.

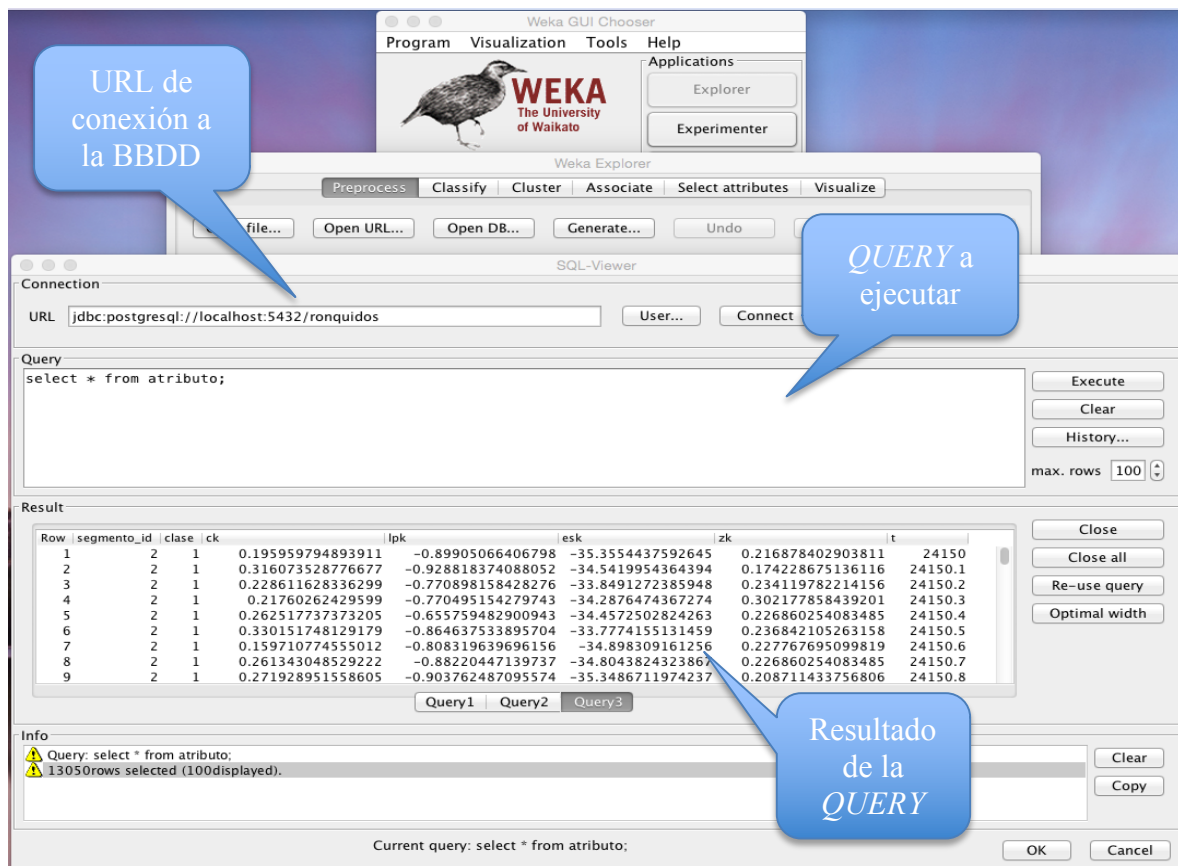


Figura 3-15: Acceso a la base de datos desde WEKA

Desde Matlab® y WEKA se pueden grabar los datos obtenidos en formato CSV para tratarlos con otras aplicaciones. Los ficheros CSV se han analizado con Microsoft-Excel®.

3.3.2 Análisis con WEKA y MATLAB®

Se procede al análisis de los datos y a la selección del clasificador utilizando WEKA y Matlab®.

3.4 Operaciones de mantenimiento de los datos

Este es un subproceso independiente que se encarga de realizar copias de todos los datos recogidos y generados.

Las copias se realizan en disco y en cinta. Como respaldo de disco, se utiliza un disco externo USB de 1TB y una unidad NAS en configuración RAID-5. Para el almacenamiento en cinta, se utiliza una librería LTO-3. En la Figura 3-17 se muestra una imagen detallada del entorno hardware utilizado.

Este subproceso queda reflejado en la Figura 3-16.

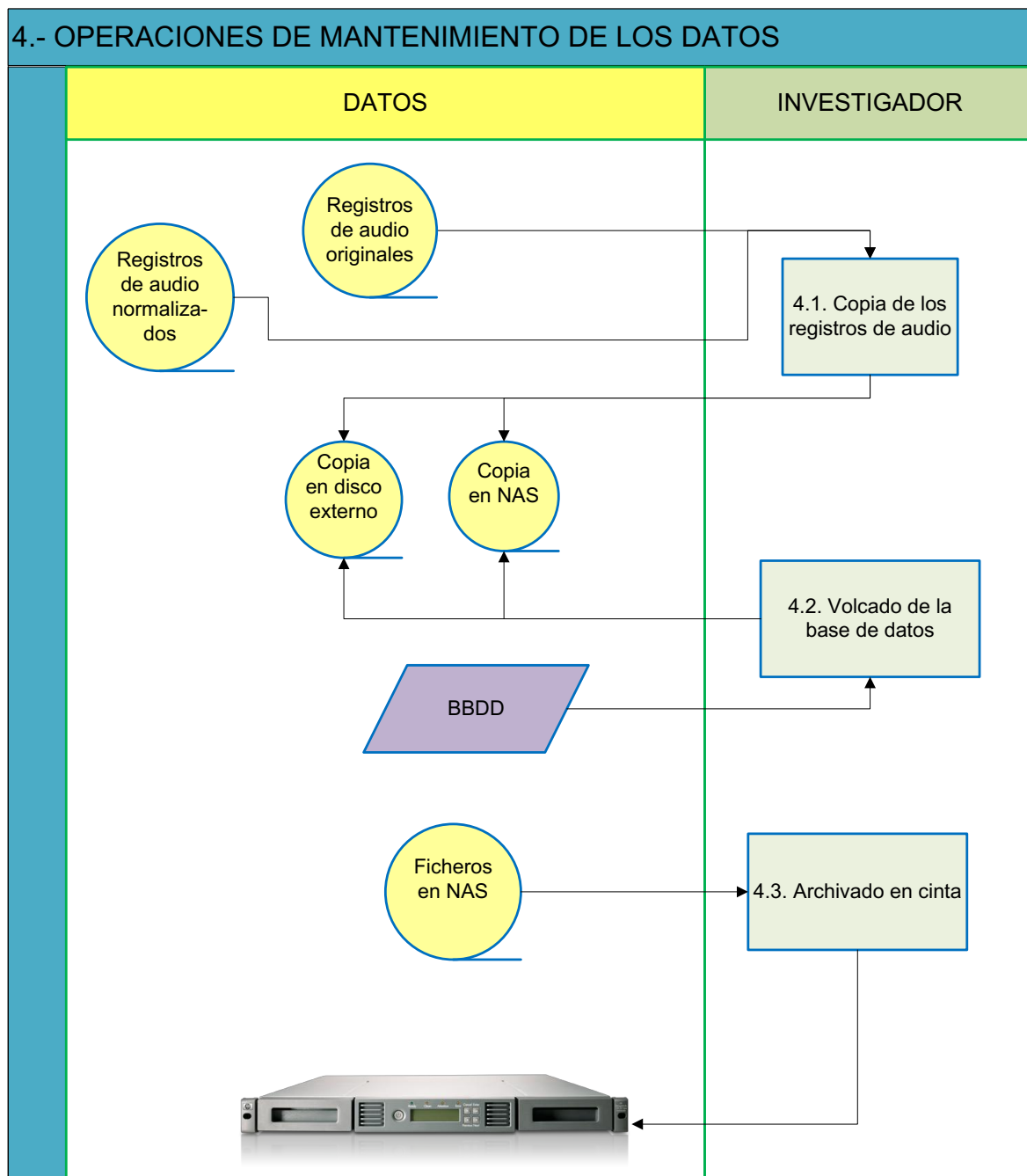


Figura 3-16: Descripción del subproceso de operaciones de mantenimiento de los datos

A continuación se detallan las tareas que componen este subproceso.

3.4.1 Copia de los registros de audio

Todos los ficheros de audio, tanto originales como procesados (normalizados), se copian en un disco externo y en un sistema NAS/RAID-5.

Los nombres de fichero son los introducidos en las tablas *“audio”* (ficheros originales) y *“datos”* (ficheros procesados).

Como se indicó en las secciones 3.1.9 y 3.2.2, los nombres de fichero son únicos.

Esta copia de seguridad se realiza cada vez que se recibe una grabación nueva o se genera un fichero de audio normalizado.

3.4.2 Volcado de la base de datos

Cada 15 días, se realiza un volcado completo de la base de datos utilizando el comando “**pg_dump**” de PostgreSQL.

El volcado se realiza sobre el disco externo, copiándose a continuación el fichero de backup generado a la unidad NAS.

De momento no se aplica ninguna política de borrado de ficheros antiguos.

En un futuro escalado del sistema, será necesario definir una política de rotación de ficheros y/o archivado de estos.

3.4.3 Archivado en cinta

Una vez al mes, se realiza una copia completa de todos los datos almacenados en la unidad NAS (ficheros de audio y volcado de la base de datos) en cartuchos de cinta LTO-3.

Los cartuchos grabados se mueven a una ubicación distinta a donde se encuentra la unidad NAS, como medida de seguridad ante posibles desastres.

De momento no se aplica ninguna política de rotación de cintas.

En un futuro escalado del sistema, será necesario definir una política de rotación de cintas y mantenimiento de las copias de seguridad.

3.5 Entorno Hardware de trabajo

En la Figura 3-17 se muestra el entorno hardware utilizado en la realización del presente trabajo.

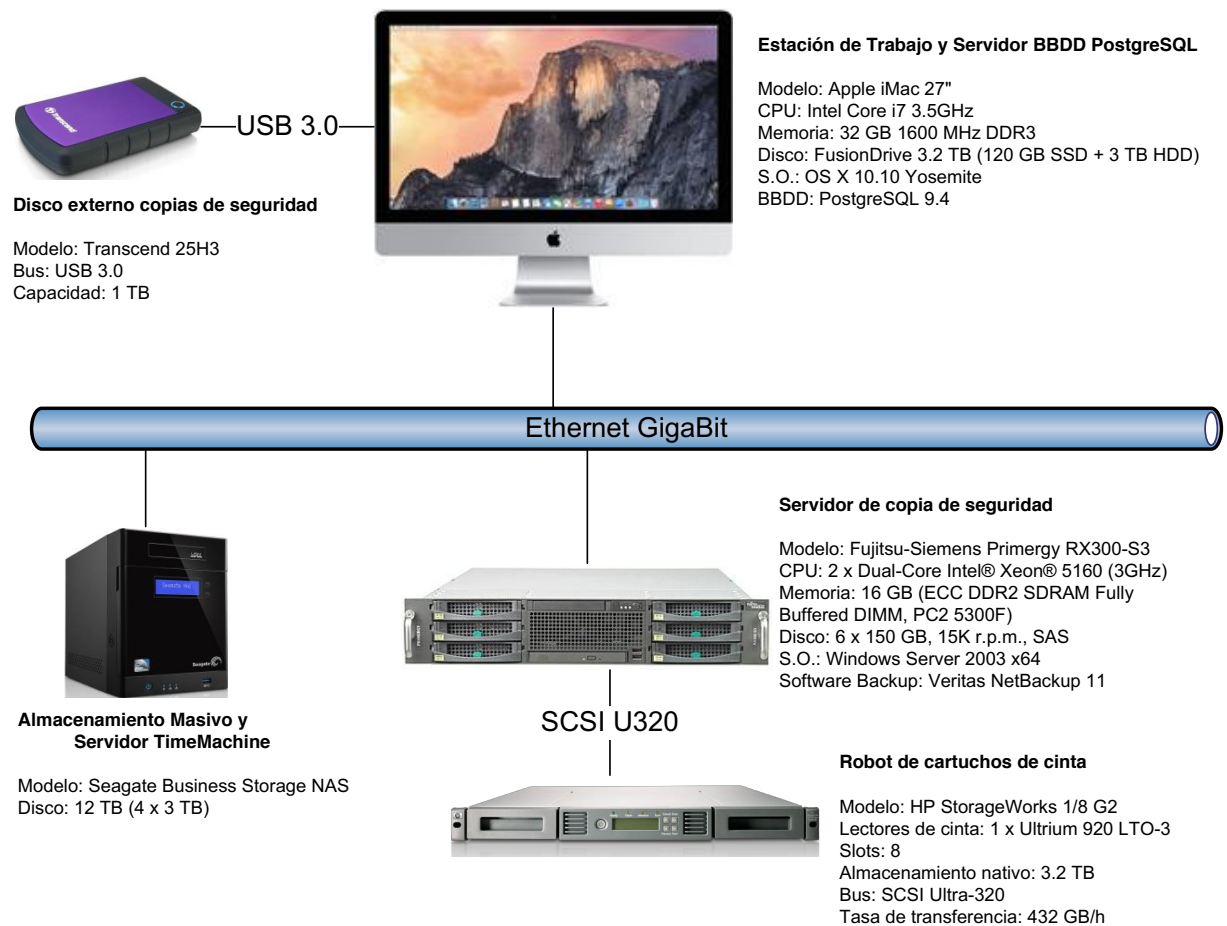


Figura 3-17: Esquema del entorno Hardware de trabajo

4 Resultados

4.1 Descripción de los datos recopilados

Se ha contado con la colaboración de 7 voluntarios que han ejercido de “*pacientes*” en la realización de este trabajo. En la tabla Tabla 4-1 se muestra los datos biométricos de los voluntarios, así como en número de grabaciones realizadas y el número de horas registradas por cada uno de ellos.

Paciente	Genero	Edad	Peso (Kg)	Altura (m)	BMI	SAHS	Nº Reg.	Horas Reg.
2	V	17	78,00	1,77	24,90	No	1	7,70
6	M	45	90,00	1,52	38,95	Si	9	55,00
8	V	6	25,00	1,30	14,79	No	3	20,56
11	V	22	91,00	1,74	30,06	No	2	9,95
13	M	24	53,00	1,67	19,00	No	3	15,27
15	V	10	59,00	1,59	23,34	No	1	7,82
17	V	57	70,00	1,62	26,67	Si	9	46,97
Total	7						28	163,27

Tabla 4-1: Datos biométricos y número de registros de los voluntarios

Se han realizado 28 grabaciones nocturnas, con un registro total de 163 horas. De estas 28 grabaciones, se han seleccionado 57 segmentos $s(n)$ que representan unos 54 minutos de grabación.

El total de ventanas $s_k(n)$ analizadas es de 32.430, de las que se han obtenido sus correspondientes vectores de atributos y clase f_k para el entrenamiento y test de los algoritmos de clasificación.

En la Tabla 4-2 y Figura 4-1 se muestra la distribución de la clasificación de los eventos respiratorios en las tres clases definidas.

Evento	Nº de Patrones
Silencio	18.002
Respiración	5.813
Ronquido	8.615

Tabla 4-2: Distribución de los eventos respiratorios clasificados

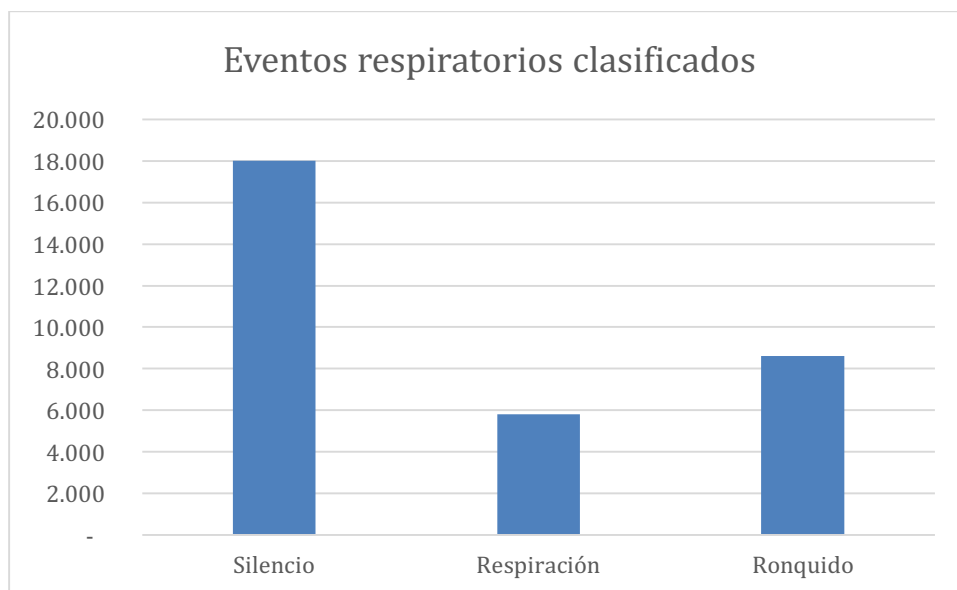


Figura 4-1: Distribución de los eventos respiratorios clasificados

De las grabadoras utilizadas (reflejadas en la Tabla 3-1), al realizar la selección de segmentos $s(n)$ su número se reduce a cuatro. La distribución de $Clase_k$ por grabadora se muestra en la Tabla 4-3 y en la Figura 4-2.

GRABADORA	SILENCIO	RESPIRACION	RONQUIDO
SONY ICD-UX200	9.678	2.874	3.528
SAMSUNG NOTE 4	4.781	1.691	3.578
APPLE IPHONE 5S	682	285	233
APPLE IPHONE 5S + RØDE SMARTLAV	2.861	963	1.276
Total	18.002	5.813	8.615

Tabla 4-3: Distribución de clases por grabadora

Del grupo de voluntarios analizado, se ha descartado uno, ya que no presentó ronquido durante las dos grabaciones que realizó.

La distribución de $Clase_k$ obtenida de cada paciente se reflejan en la Tabla 4-4 y en la Figura 4-3.

La primera columna de la Tabla 4-4 corresponde con el identificador único asignado por la BBDD a cada registro de paciente. Se observa que aunque solo hay siete pacientes, los identificadores llegan hasta 13, y no son consecutivos en incremento de 1. Esto es debido al sistema de secuencia concurrente implementado por PostgreSQL: para evitar bloqueos en transacciones concurrentes que obtienen números de la misma secuencia, una operación de obtener número no se puede retroceder (*roll-back*), por lo que las transacciones que son canceladas dejan “huecos” no usados en la secuencia de valores asignados.

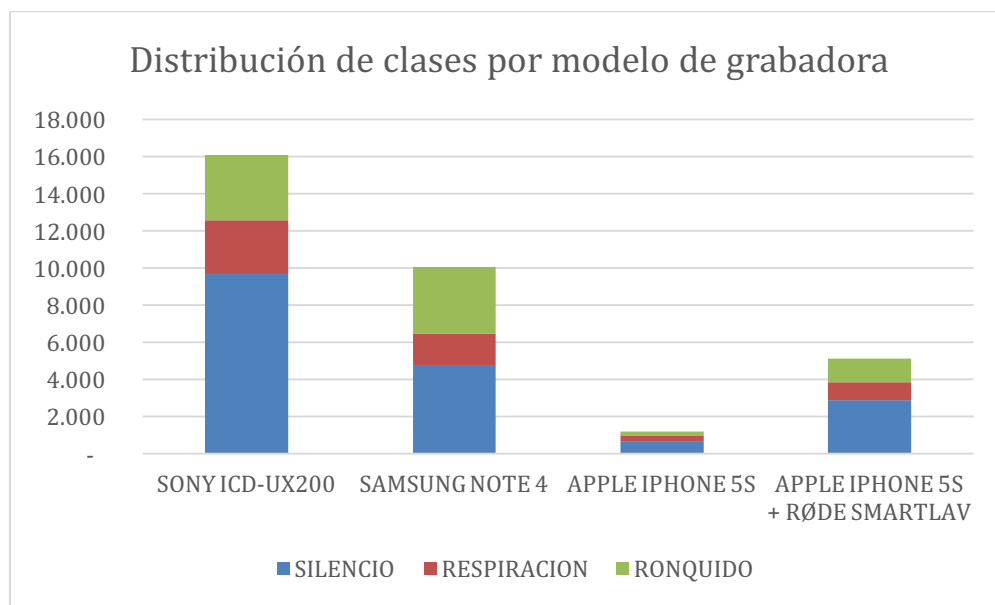


Figura 4-2: Distribución de clases por modelo de grabadora

PACIENTE	SILENCIO	RESPIRACION	RONQUIDO
13	1.703	493	404
15	1.766	407	527
17	4.781	1.691	3.578
2	3.820	748	2.062
6	3.543	1.248	1.509
8	2.389	1.226	535
Total	18.002	5.813	8.615

Tabla 4-4: Distribución de clases por paciente

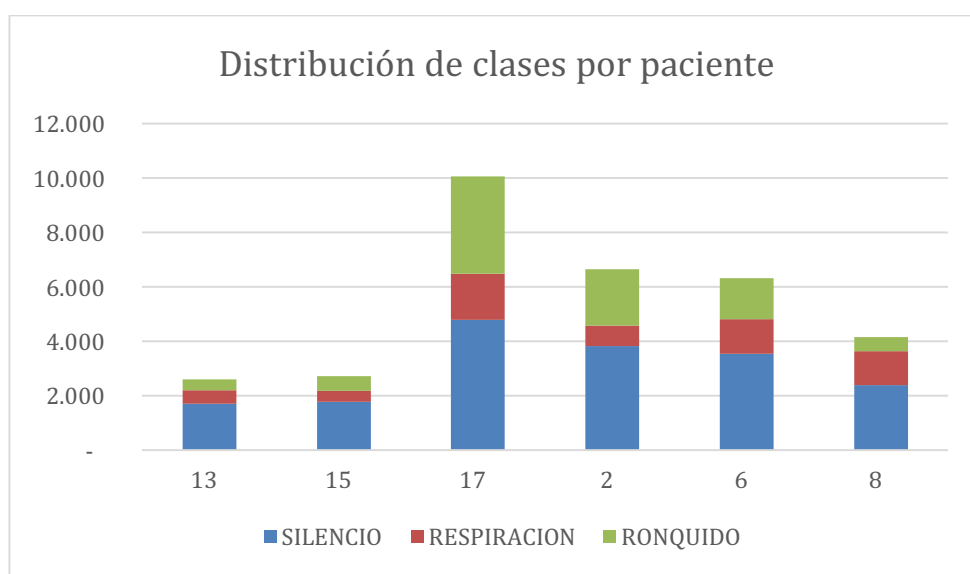


Figura 4-3: Distribución de clases por paciente

El grupo final se compone de 2 mujeres y 4 varones. La distribución de $Clase_k$ por género se puede muestra en la Tabla 4-5 y en la Figura 4-4.

GENERO	SILENCIO	RESPIRACION	RONQUIDO
VARON	12.756	4.072	6.702
MUJER	5.246	1.741	1.913
Total	18.002	5.813	8.615

Tabla 4-5: Distribución de clases por género

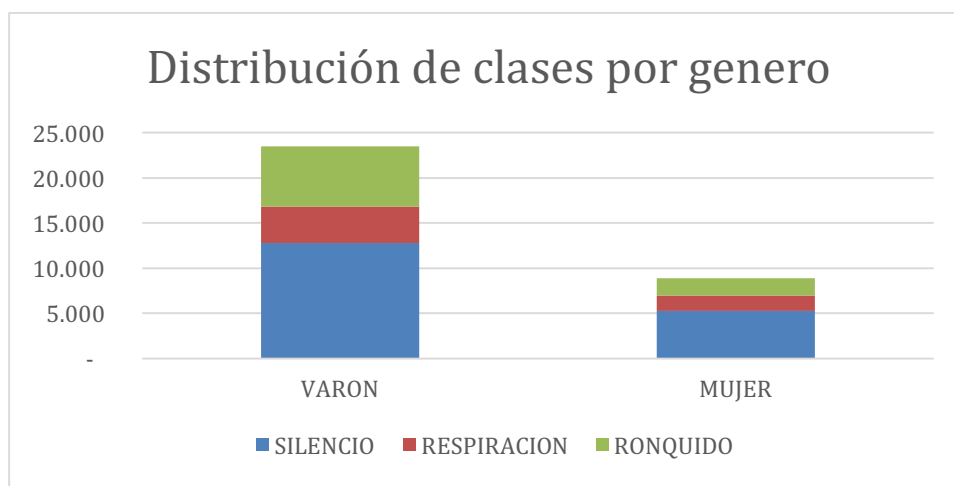


Figura 4-4: Distribución de clases por género

Las edades de los voluntarios están comprendidas entre los 6 y los 57 años. La distribución de $Clase_k$ por grupo de edad se muestra en la Tabla 4-6 y en la Figura 4-5 .

EDAD	SILENCIO	RESPIRACION	RONQUIDO
0 - 15	4.155	1.633	1.062
16 - 25	5.523	1.241	2.466
26 - 45	-	-	-
> 45	8.324	2.939	5.087
Total	18.002	5.813	8.615

Tabla 4-6: Distribución de clases por edad

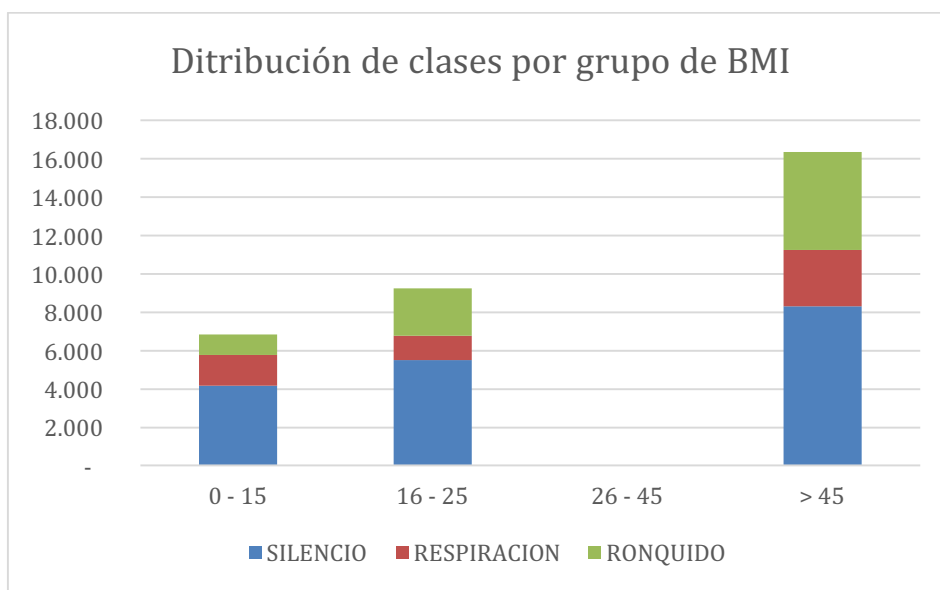


Figura 4-5: Distribución de clases por grupo de edad

Entre los voluntarios, hay una mujer y un varón con SAHS diagnosticados mediante PSG. A la segunda mujer se le descartó SAHS mediante la realización de una prueba PSG. Al resto de varones no se les ha realizado la prueba de PSG.

La distribución de $Clase_k$ por diagnóstico de SAHS se muestra en la Tabla 4-7 y en la Figura 4-6.

SAHS DIAG.	SILENCIO	RESPIRACION	RONQUIDO
SI	8.324	2.939	5.087
NO	9.678	2.874	3.528
Total	18.002	5.813	8.615

Tabla 4-7: Distribución de clases por SAHS diagnosticado

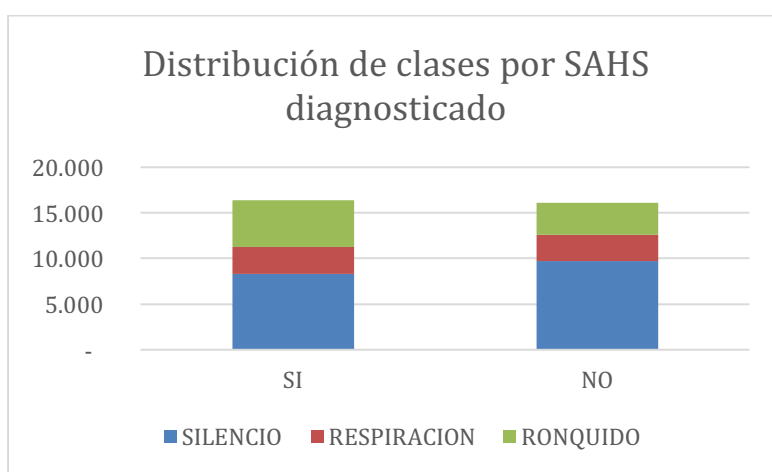


Figura 4-6: Distribución de clases por SAHS diagnosticado

El índice BMI de los voluntarios está comprendido entre 14,79 y 38,95. La Organización Mundial de la Salud define que a un adulto con un BMI entre 25 y 29,9 se le considera con sobrepeso, y con un valor superior a 30 como obeso en distintos grados (World Health Organization (WHO), 2015).

La distribución de $Clase_k$ por grupo de BMI (en la sección 3.1.7 se describe el método de cálculo de este índice) se puede muestra en la Tabla 4-8 y en la Figura 4-7.

BMI	SILENCIO	RESPIRACION	RONQUIDO
NORMAL (BMI<25)	9.678	2.874	3.528
SOBREPESO (25≤BMI≤29,9)	4.781	1.691	3.578
OBESO (BMI≥30)	3.543	1.248	1.509
Total	18.002	5.813	8.615

Tabla 4-8: Distribución de clases por grupo de BMI

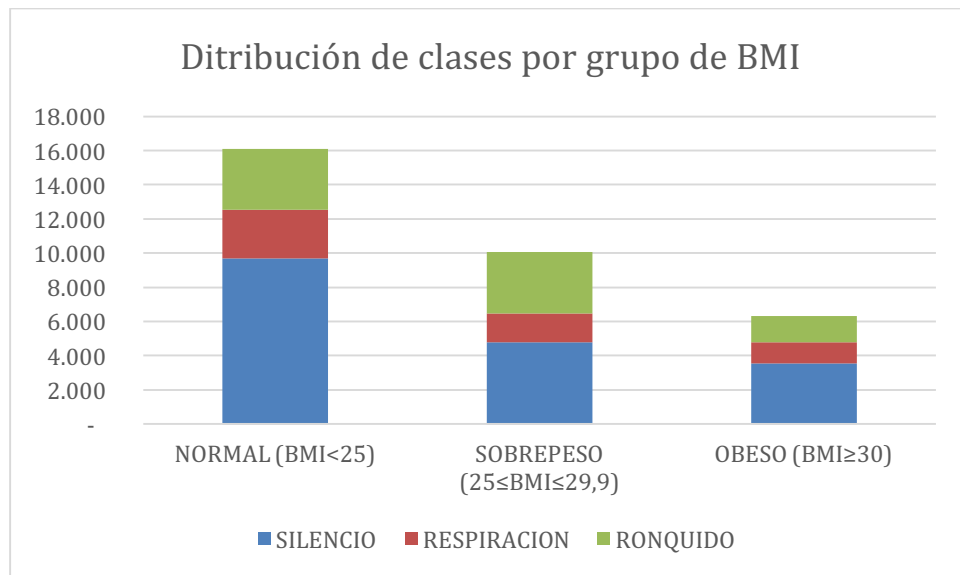


Figura 4-7: Distribución de clases por grupo de BMI

Esta distribución por grupo de BMI puede estar algo distorsionada, pues en el grupo de voluntarios hay dos niños, y los valores de BMI están referidos a personas adultas.

4.2 Clasificación automática de los eventos respiratorios en base a las características del audio

Un primer análisis de los cuatro atributos de audio calculados consiste en visualizar la distribución de los valores de f_k (vector de atributos y clase). Se calcula el histograma de distribución de valores de cada atributo para cada clase de evento respiratorio.

En Figura 4-8, Figura 4-9, Figura 4-10 y Figura 4-11 se recoge la distribución de valores para cada atributo, una curva por cada clase de evento respiratorio. Esta visualización nos

permite en una primera aproximación evaluar los difícil o sencillo que va a resultar la clasificación de los eventos, observando la separación de las distribuciones de los valores. Cuanto mayor sea la separación de las distribuciones de las clases, más sencilla será su clasificación.

Analizando la distribución de Z_k (*Zero Crossing Rate*, Figura 4-8), se observa que la distribución de silencio coincide con la distribución de respiración en valores bajos, lo que dificultaría su separación en el modelo. Este es un hecho esperado, ya que la respiración se asemeja bastante al ruido blanco de fondo.

Así mismo, la distribución de la clase *silencio* presenta un aspecto bimodal. Se podría sospechar que parte de los valores del máximo más alto (valores en torno a 0,25-0,30) podrían estar mal clasificados y corresponder a la clase *respiración*. Esto puede ser debido al problema de no tener establecido un criterio claro a la hora de clasificar manualmente la clase de las ventanas $s_k(n)$, tal y como se indicó en la sección 3.2.5.

Respecto a los valores de Z_k de los eventos *ronquido*, se observa que aunque presentan una distribución muy amplia, tiene su máximo en torno a 0,1. Esto más o menos coincide con lo indicado en la sección 3.2.6.1 de que si Z_k es bajo, la señal presenta una estructura similar a la locución (Bachu et al., 2010; Karunajeewa et al., 2008). Se podría argumentar que esto no es válido porque los eventos de *silencio* también presentan máximos en su distribución a valores bajos (en torno a 0,18). Una explicación es que los pasajes de silencio contienen muchos elementos con estructura en el ruido de fondo: pájaros cantando, vehículos que pasan por la calle, tic-tac de un reloj, pisadas de personas caminado por la casa, conversaciones de fondo, etc. Pero por otro lado, los eventos de *respiración* presentan su máximo a valores más altos, lo que apoya esta hipótesis, ya que los sonidos de inspiración/expiración se podrían asimilar a un ruido rosa. Todo esto será necesario tenerlo en cuenta en futuros estudios.

Respecto a la distribución de valores de Es_k (*Short Time Energy*, Figura 4-9), se observa que presenta los mismos problemas en los eventos de *silencio* que Z_k respecto a los ruidos de fondo. Hay que recordar, tal y como se indicó en la sección 3.2.6.2, que esta medida es muy dependiente de los parámetros de grabación, calidad del equipo, sensibilidad del micrófono y distancia a la fuente sonora.

Los *silencios* deberían presentar los valores más bajos de energía, cosa que ocurre con los valores agrupados por debajo de -55 dB. Los valores que aparecen en torno a los -35 dB podrían corresponder a los ruidos de fondo indicados, ya que en algunas grabaciones el nivel de la señal cuando se producían era muy alto.

Los eventos de *respiración* también deberían presentar valores bajos de energía. Esto más o menos se cumple, ya que los valores se agrupan por debajo de -50 dB, pero presenta una agrupación de valores importante entre -35 y -30 dB. Este efecto puede explicarse por las diferentes características de los equipos de registro utilizados y por el entorno de grabación (posición del micrófono y distancia a la fuente sonora).

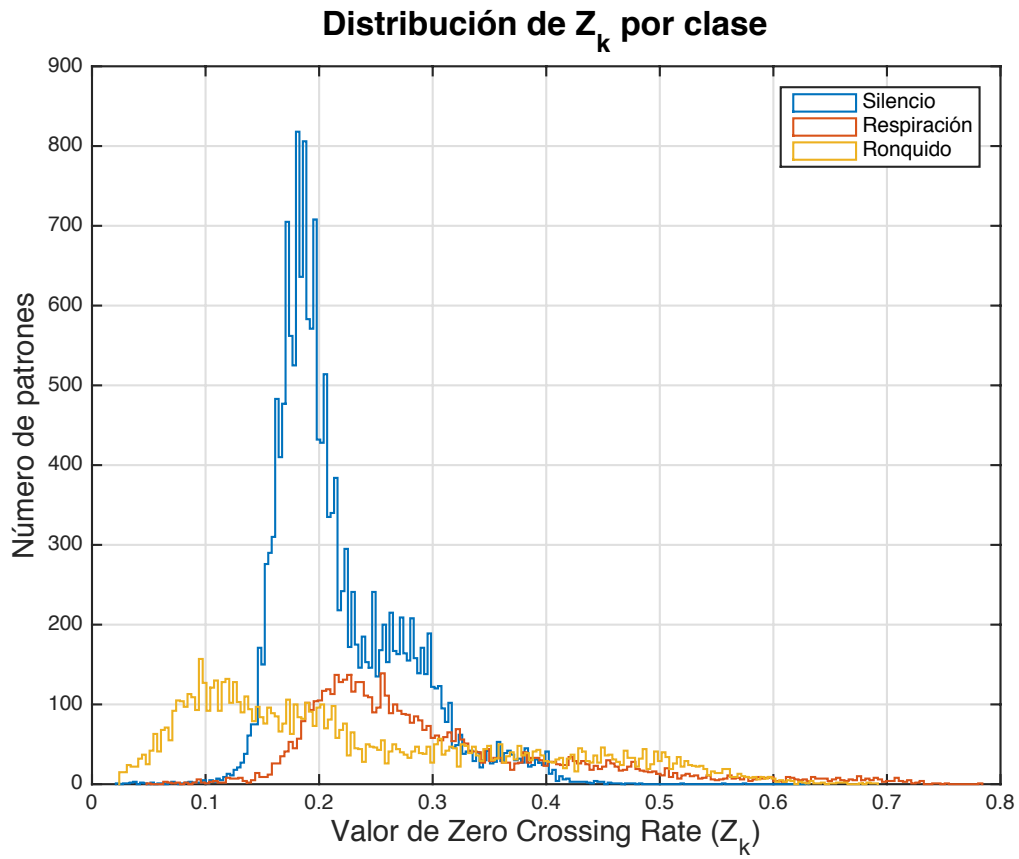


Figura 4-8: Distribución de Z_k en función de la clase de evento respiratorio

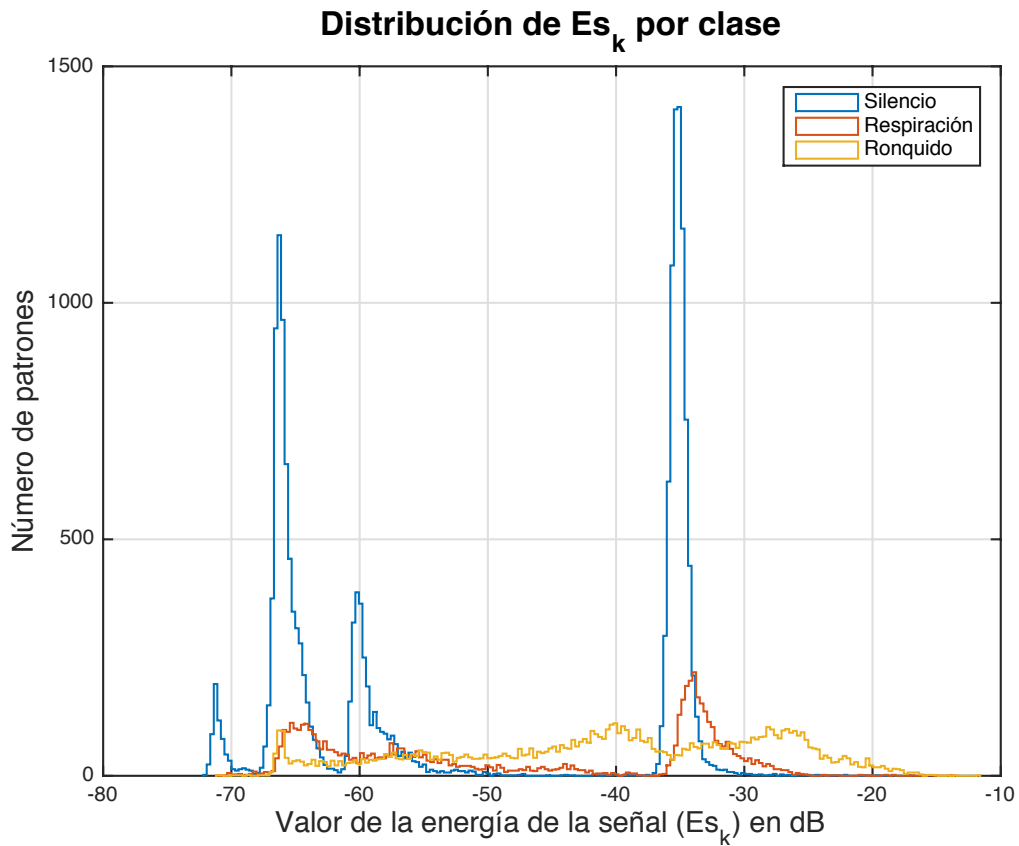


Figura 4-9: Distribución de Es_k en función de la clase de evento respiratorio

Los valores de Es_k para los eventos *ronquido* deberían ser los de mayor energía, como se indicó en la sección 3.2.6.2. Se observa que son los que presentan mayores niveles de energía (hay un pico entre -30 dB y -20 dB), pero presentan una distribución muy amplia, entre -60 dB y -20 dB. De nuevo este efecto puede deberse a las diferentes características de los equipos de registro y entorno de grabación.

De todo lo anterior, se desprende la necesidad de realizar algún tipo de normalización de los valores de Es_k en función del equipo de registro y del entorno de grabación. Quizá se podría utilizar la primera media hora de grabación para definir el “cero” del registro, o como base para un método de filtrado. Todo esto será necesario tenerlo en cuenta en futuras investigaciones.

Sobre la distribución de los valores de C_k (*Coficiente de Correlación a 1 ms de retardo*, Figura 4-10), se debería esperar que los pasajes con estructura (voz/ronquido) presentasen una correlación más alta entre muestras consecutivas que las secciones sin estructura (ver sección 3.2.6.3). Se observa que parte de los valores de *ronquido* presentan la correlación más alta (mayor de 0,4 en valor absoluto), pero su distribución de valores se solapa bastante con las otras dos clases.

Las clases *respiración* y *silencio* se solapan totalmente, por lo que esta medida, a priori, no sería buena para realizar su clasificación, pues presentaría poca discriminación.

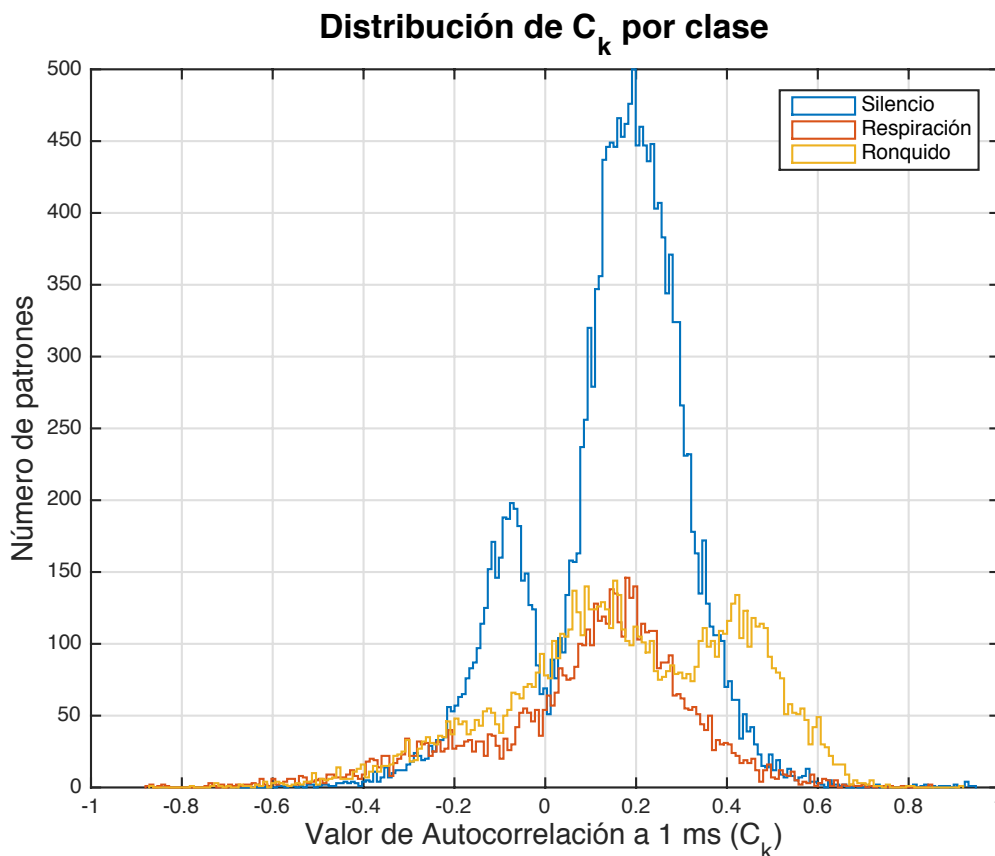


Figura 4-10: Distribución de C_k en función de la clase de evento respiratorio

De la distribución de valores de Lp_k (*Primer Coeficiente Predictor de Análisis LPC de 12 polos*, Figura 4-11) se observa que las tres clases presentan un solapamiento muy alto. En este caso, las zonas de mayos discriminación corresponderían a valores por debajo de -1.25 o superiores a 0, donde sería clara la separación de la clase *ronquido*.

Según se indicó en la sección 3.2.6.4, era de esperar que los valores de Lp_k fueran considerablemente diferentes para cada clase, pero no se aprecia claramente.

Las distribuciones para *respiración* y *ronquido* parecen multimodales, mientras que la distribución de *silencio* es multimodal. Esto último quizá se pueda explicar por los ruidos de fondo con estructura que aparecen en los pasajes de silencio.

Aunque por separado, parece que las características seleccionada no han de presentar una buena separación en la clasificación, hay que analizar cómo se comportan cuando se manejan en conjunto. Esto se analiza más adelante en esta sección.

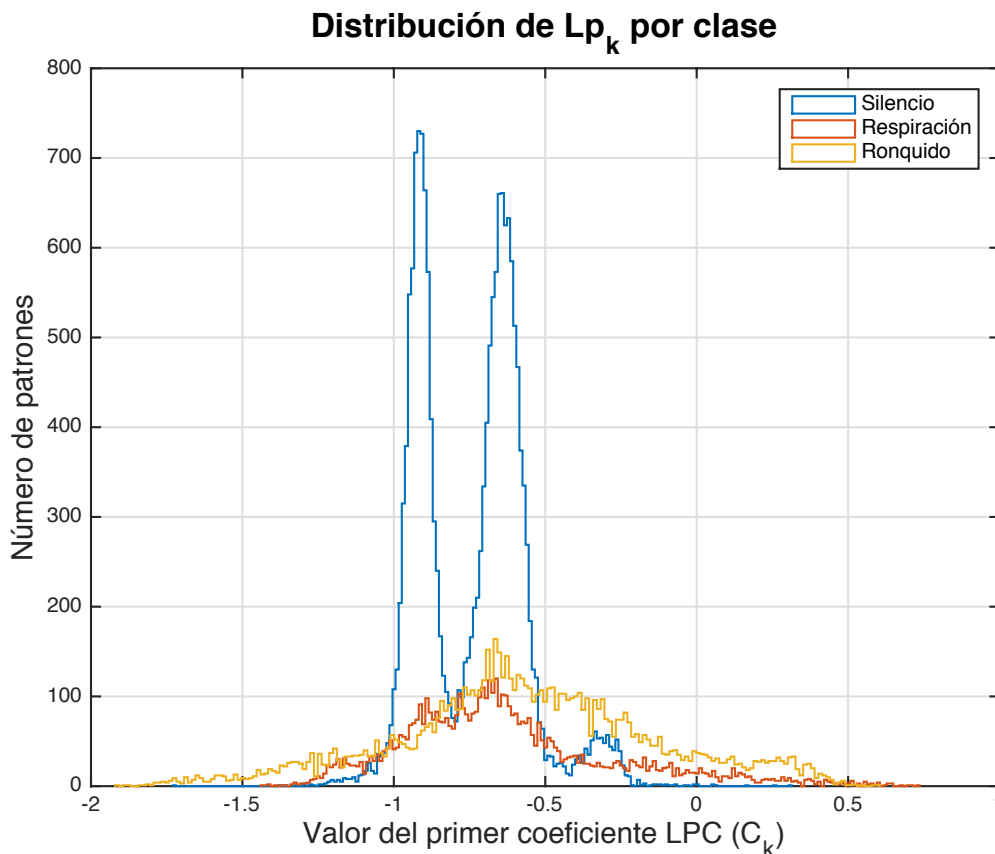


Figura 4-11: Distribución de Lp_k en función de la clase de evento respiratorio

Podemos representar de forma gráfica las relaciones entre los atributos dos a dos para observar su distribución en el plano y ver posibles relaciones. En la Figura 4-12 observamos estas relaciones para los cuatro atributos. El color del punto indica la clase a la que pertenece: Rojo/Silencio, Azul/respiración, Verde/Ronquido. El pie de cada subfigura indica la característica asociada a cada eje: primer valor, eje x; segundo valor, eje y.

Observamos que los componentes presentan muy baja correlación entre ellos, excepto en el caso de Lp_k vs. Z_k , que vemos que presenta cierta tendencia a agrupar los valores en la diagonal.

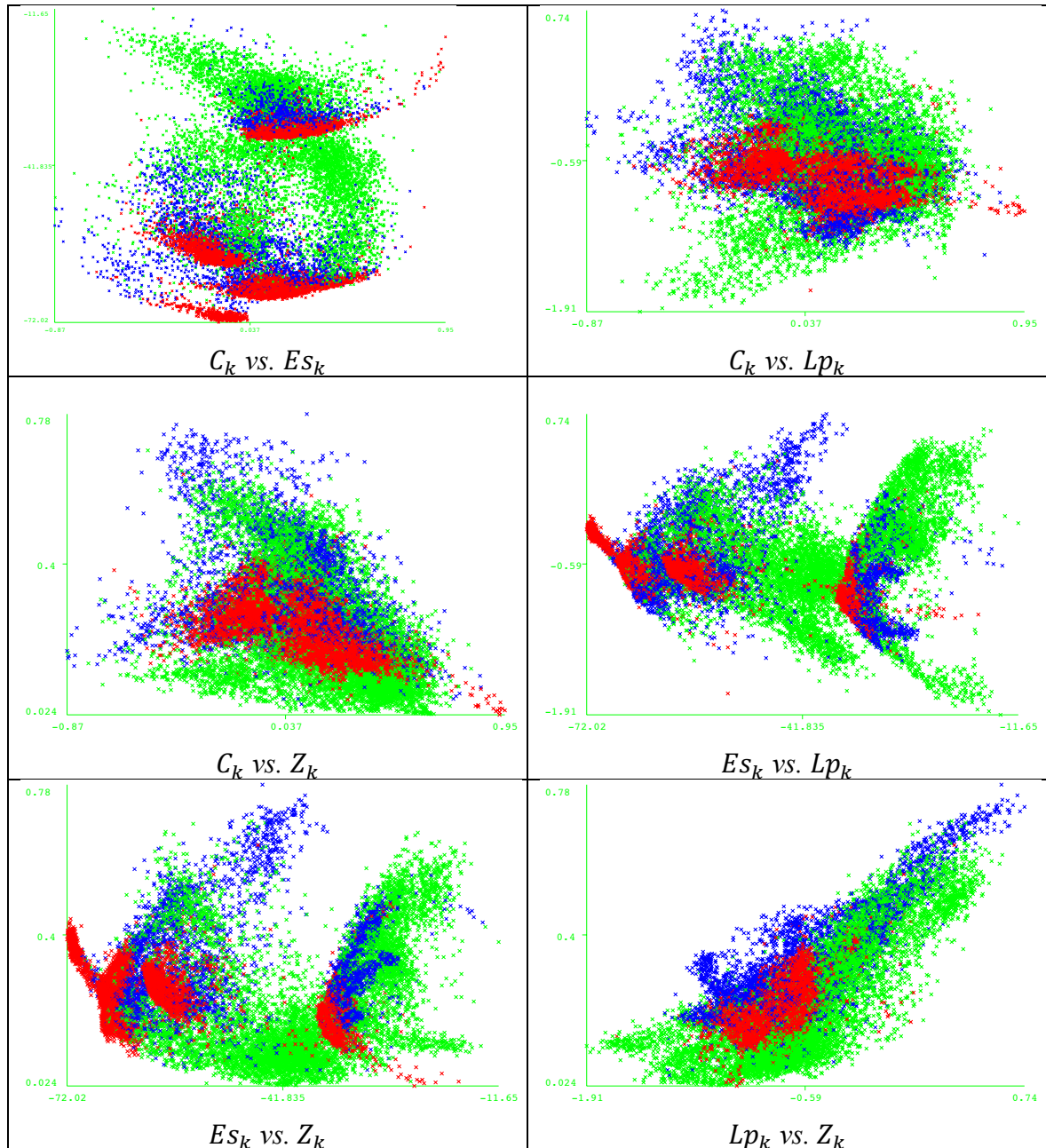


Figura 4-12: Relaciones entre los atributos de f_k para los eventos respiratorios

Para comprobar la independencia de los componentes del vector f_k , se calcula la matriz de correlación de los cuatro atributos que lo componen. En la Tabla 4-9 se recogen los valores del cálculo de la matriz.

	C_k	Es_k	Lp_k	Z_k
C_k	1	0.19	-0.23	-0.47
Es_k	0.19	1	-0.18	-0.04
Lp_k	-0.23	-0.18	1	0.77
Z_k	-0.47	-0.04	0.77	1

Tabla 4-9: Matriz de correlación de los atributos de f_k

Observamos que no hay correlación alta entre los componentes, y que los que presentan mayor correlación son Lp_k vs. Z_k , tal y como se había observado en la Figura 4-12. Esta mayor correlación entre Lp_k y Z_k se puede explicar porque ambos atributos representan en mayor o menor medida el espacio espectral de $s_k(n)$. En futuros estudios habrá que realizar un análisis más detallado para verificar si se puede eliminar alguna de las dos características, simplificando el cálculo *on-line*.

Ya que los componentes presentan baja correlación entre ellos, no descartaremos ninguno y se utilizarán los cuatro para la clasificación de los eventos respiratorios.

Las pruebas de los distintos clasificadores se han realizado con WEKA. Se ha dividido el 66% de los patrones para entrenamiento (21404 patrones) y el 34% para test (11026 patrones). Las descripciones de los clasificadores utilizados y sus parámetros se recogen en el anexo A.

Los resultados que se recogen son los siguientes:

- **% Clasif. Correctos:** Porcentaje de instancias de test correctamente clasificadas
- **True Positive Rate:** tasa que refleja la proporción de instancias que fueron clasificadas como clase x, entre todas las instancias que verdaderamente tienen clase x. Representa el parámetro de Sensibilidad.
- **False Positive Rate:** tasa que refleja la proporción de ejemplos que fueron clasificados como clase x, pero pertenece a una clase diferente, entre todos los ejemplos que no son de clase x. Representa 1-Especificidad.
- **ROC Área:** Área bajo la curva ROC o AUC. A modo de guía para interpretar las curvas ROC se han establecido los siguientes intervalos para los valores de AUC:
 - [0.5, 0.6]: Test malo.
 - [0.6, 0.75]: Test regular.
 - [0.75, 0.9]: Test bueno.
 - [0.9, 0.97]: Test muy bueno.
 - [0.97, 1]: Test excelente.

A modo de recordatorio, indicar que la Sensibilidad indica la capacidad del modelo de clasificación para dar como casos positivos los casos realmente positivos: la proporción de enfermos, o eventos respiratorios, correctamente identificados. Se define como:

$$Sensibilidad = \frac{VP}{VP + FN} \quad (8)$$

Donde VP son los Verdaderos Positivos y FN los Falsos Negativos.

Por otro lado, la Especificidad indica la capacidad de nuestro modelo de clasificación para dar como casos negativos los casos que son realmente negativos: proporción de pacientes sanos correctamente clasificados. Se define como:

$$\text{Especificidad} = \frac{VN}{VN + FP} \quad (9)$$

Donde VN son los Verdaderos Negativos y FP los Falsos Positivos.

True Positive Rate, *False Positive Rate* y *ROC Área* se desglosan en sus valores para las clases de *silencio*, *respiración* y *ronquido*, y para el promedio ponderado.

En la Tabla 4-10 se recogen los resultados de los distintos modelos de clasificación utilizados, así como los parámetros de su configuración. El significado de los parámetros de los distintos clasificadores utilizados se pueden consultar en el anexo A.

Atributos seleccionados: C_k Lp_k Z_k Es_k $Clase_k$

Clase Clasif.	Parám. Clasif.	% Clasif. Correctos	True Positive Rate (Sensibilidad)				False Positive Rate (1 – Especificidad)				ROC Área (AUC)			
			Sil.	Res.	Ron.	W.A.	Sil.	Res.	Ron.	W.A.	Sil.	Res.	Ron.	W.A.
ZeroR	n.a.	55.7954	1	0	0	0.558	1	0	0	0.558	0.5	0.5	0.5	0.5
NaiveBayes Simple	n.a	69.4812	0.931	0.181	0.542	0.695	0.476	0.068	0.053	0.292	0.801	0.6747	0.853	0.787
Logistic	-R 1.0E-8 -M -1	72.5104	0.973	0.26	0.516	0.725	0.49	0.031	0.045	0.291	0.748	0.747	0.747	0.748
Multilayer Perceptron 1 capa oculta 3 neuronas	-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H 3	80.3465	0.96	0.442	0.717	0.803	0.28	0.061	0.031	0.175	0.902	0.782	0.916	0.792
Multilayer Perceptron 1 capa oculta 4 neuronas	-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H 4	80.4644	0.947	0.42	0.764	0.805	0.238	0.059	0.057	0.158	0.917	0.733	0.917	0.884
Multilayer Perceptron 1 capa oculta 5 neuronas	-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H 5	81.9064	0.951	0.462	0.783	0.819	0.23	0.059	0.042	0.15	0.924	0.822	0.932	0.908
Multilayer Perceptron 1 capa oculta 6 neuronas	-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H 6	81.9336	0.928	0.558	0.768	0.819	0.192	0.083	0.037	0.132	0.927	0.813	0.927	0.906
Multilayer Perceptron 1 capa oculta 8 neuronas	-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H 8	83.0038	0.935	0.568	0.784	0.83	0.182	0.072	0.041	0.125	0.934	0.847	0.941	0.92

Clase Clasif.	Parám. Clasif.	% Clasif. Correctos	True Positive Rate (Sensibilidad)				False Positive Rate (1 – Especificidad)				ROC Área (AUC)			
			Sil.	Res.	Ron.	W.A.	Sil.	Res.	Ron.	W.A.	Sil.	Res.	Ron.	W.A.
SMO (SVM) Kernel Lineal	-C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K PolyKernel -C 250007 -E 1.0	72.9458	0.98	0.279	0.504	0.729	0.495	0.031	0.036	0.291	0.744	0.672	0.776	0.74
SMO (SVM) Kernel Lineal Low Order	-C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K PolyKernel -C 250007 -E 1.0 -L	72.9276	0.98	0.279	0.504	0.729	0.495	0.031	0.036	0.291	0.744	0.672	0.776	0.697
SMO (SVM) Kernel Cuadrático	-C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K PolyKernel -C 250007 -E 2.0	75.984	0.975	0.361	0.574	0.76	0.421	0.042	0.027	0.249	0.779	0.744	0.802	0.779
SMO (SVM) Kernel Cuadrático Low Order	-C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K PolyKernel -C 250007 -E 2.0 -L	78.3421	0.976	0.346	0.673	0.783	0.355	0.044	0.033	0.214	0.812	0.737	0.844	0.807
J48	-C 0.25 -M 2	83.8382	0.927	0.594	0.817	0.838	0.158	0.069	0.047	0.113	0.931	0.84	0.925	0.913
Random Tree	-K 0 -M 1.0 -S 1	79.3397	0.87	0.553	0.795	0.793	0.165	0.097	0.073	0.128	0.853	0.728	0.861	0.832
Random Forest 100 árboles	-I 100 -K 0 -S 1	85.2621	0.939	0.611	0.833	0.853	0.157	0.06	0.039	0.108	0.952	0.898	0.958	0.944

Tabla 4-10: Resultados de los distintos clasificadores para los eventos respiratorios

Como se observa en la tabla, los clasificadores que mejor funcionan son los árboles J48, RandomForest, y Perceptrón Multicapa con una capa oculta de 8 neuronas, todos con una AUC por encima de 0.91.

También se observa que todos los clasificadores tienen peor desempeño a la hora de clasificar la respiración, pues como se indicó en la sección 3.2.5, son los eventos que más dificultades presentaron a la hora de clasificarlos manualmente. La respiración, al presentar frontera con el silencio (o ruido de fondo) y el ronquido simultáneamente, presenta el doble de incertidumbre que los otros dos eventos respiratorios.

No se ha utilizado SMO (SVM) con kernel RBF debido a dos factores:

1. Implementación en el dispositivo móvil: el uso de exponenciales penalizaría mucho el tiempo de ejecución de la clasificación.
2. Para su optimización hay que realizar una búsqueda paramétrica exhaustiva de γ .

La red neuronal utilizada en el perceptrón multicapa queda reflejada en la Figura 4-13. En los distintos modelos probados, se ha modificado el número de neuronas de la capa oculta (círculos rojos).

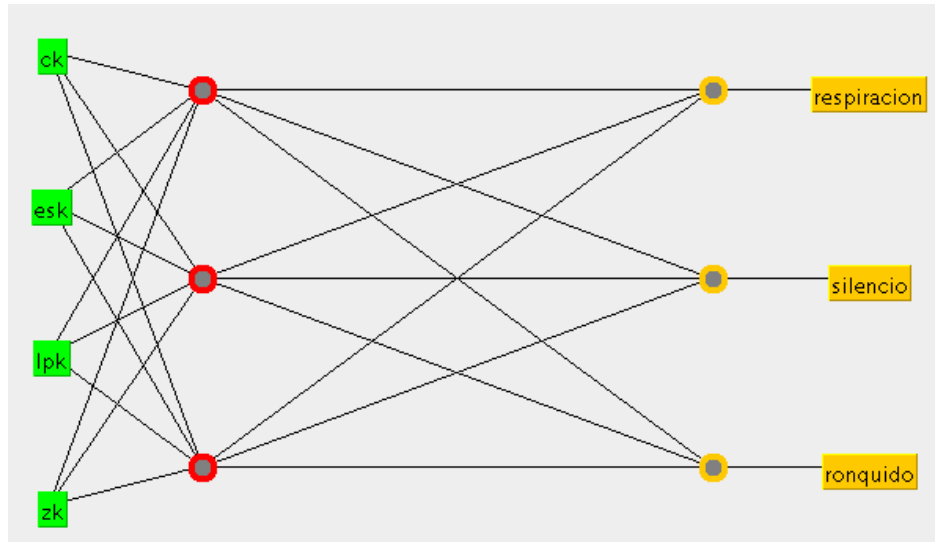


Figura 4-13: Red neuronal del perceptrón de una capa oculta con tres neuronas

El siguiente paso de análisis consiste en añadir atributos recogidos en el formulario STOP-BANG, para ver si estos factores ayudan a mejorar la clasificación de los eventos respiratorios.

De los atributos obtenidos en la sección 3.1.7, se descartan los siguientes:

- Fecha de nacimiento: Se utiliza la edad
- Perímetro del cuello. Datos incompletos
- Comentarios del paciente y del investigador: no aplican
- BMI ≥ 35 : Usamos BMI directamente
- Perímetro cuello ≥ 40 : Datos incompletos
- Edad ≥ 50 : Usamos edad directamente
- Ronca: todos roncan
- Se levanta cansado: datos incompletos
- Es hipertenso: todos son no-hipertensos
- Apneas observadas: datos incompletos
- SAHS diagnosticado: es lo que buscamos en último término

Con lo que los atributos a estudiar son:

- Altura (Figura 4-14)
- Peso (Figura 4-15)

- Edad (Figura 4-16)
- BMI (Figura 4-17)
- Género (Figura 4-18)

Aunque el índice BMI es una combinación de peso y altura, como se mostró en la ecuación (1) de la sección 3.1.7, se mantienen los tres para comprobar su aportación por separado.

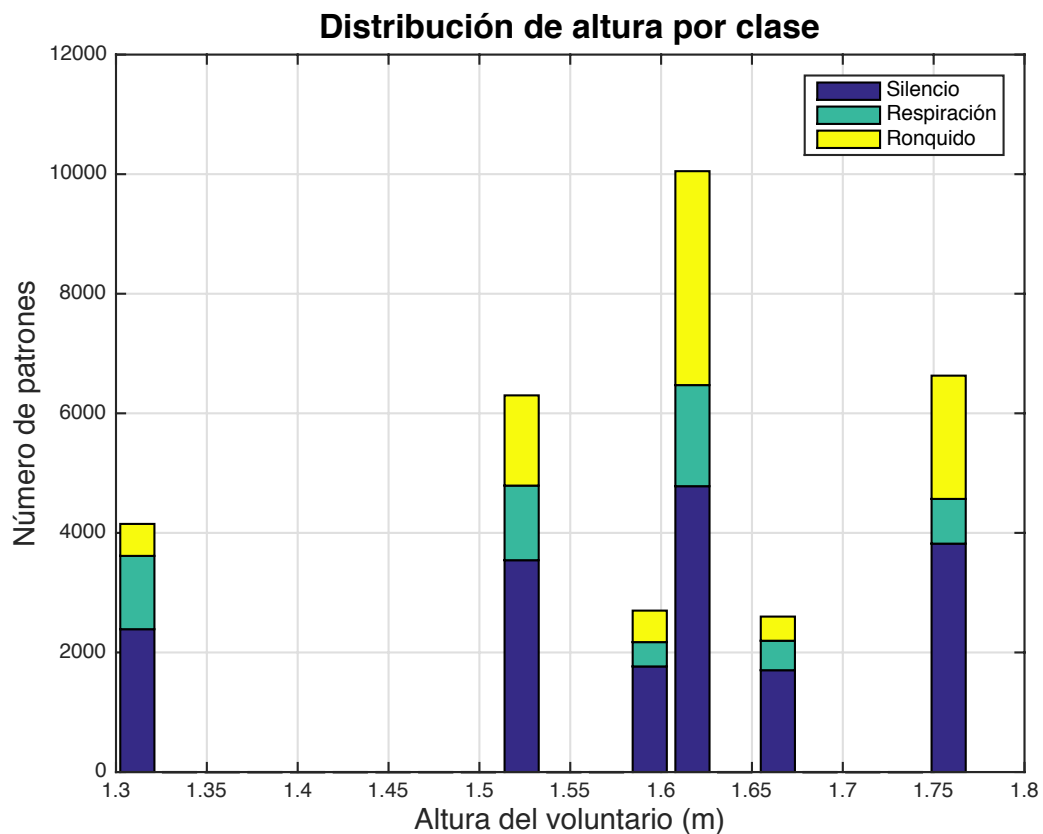


Figura 4-14: Distribución de la altura de los voluntarios por clase de evento respiratorio

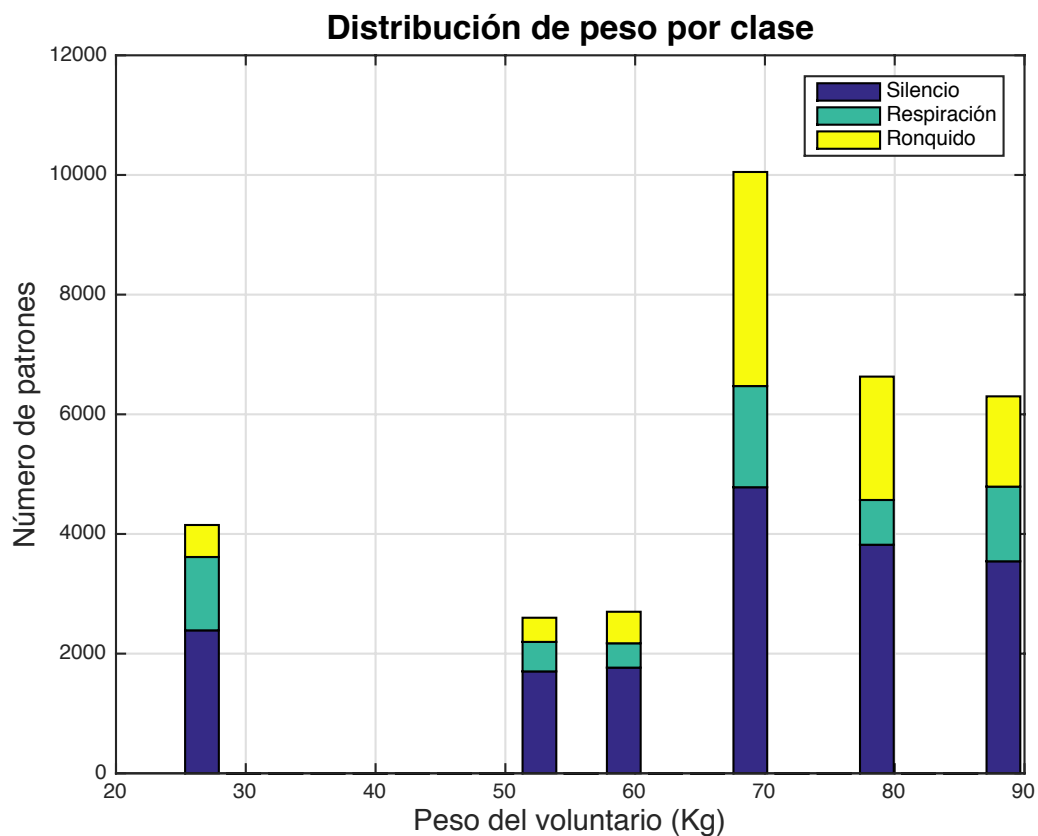


Figura 4-15: Distribución del peso de los voluntarios por clase de evento respiratorio

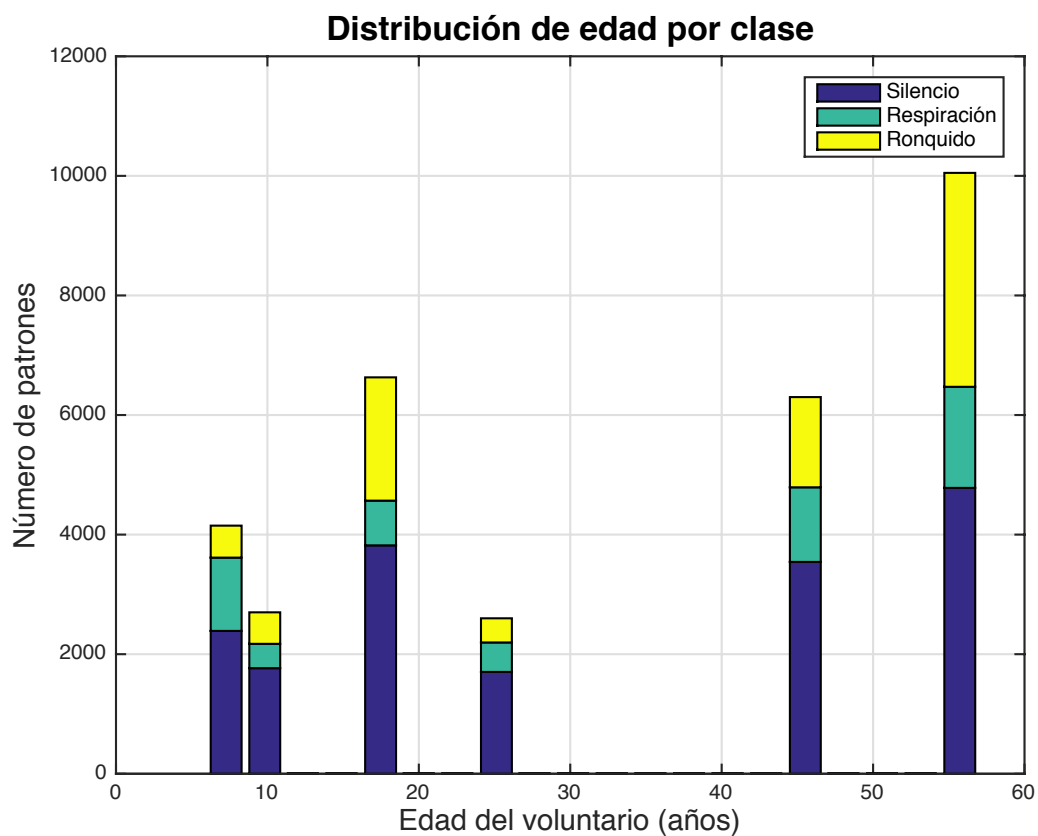


Figura 4-16: Distribución de la edad de los voluntarios por clase de evento respiratorio

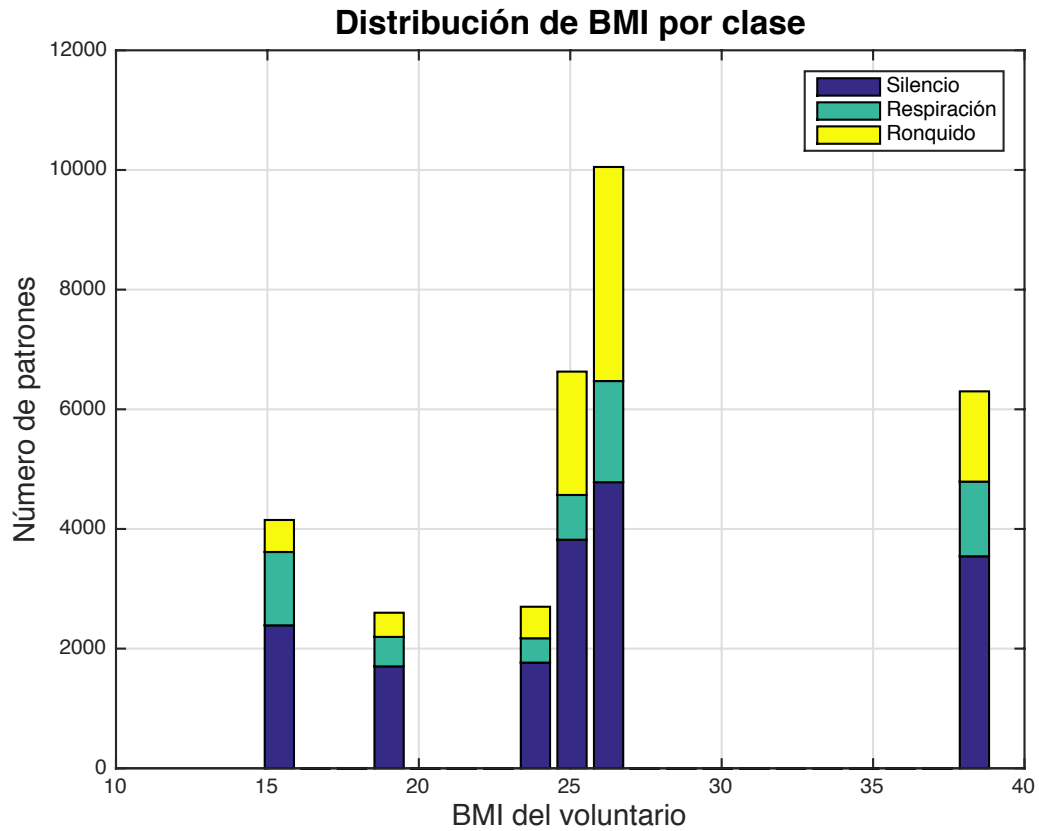


Figura 4-17: Distribución de BMI de los voluntarios por clase de evento respiratorio

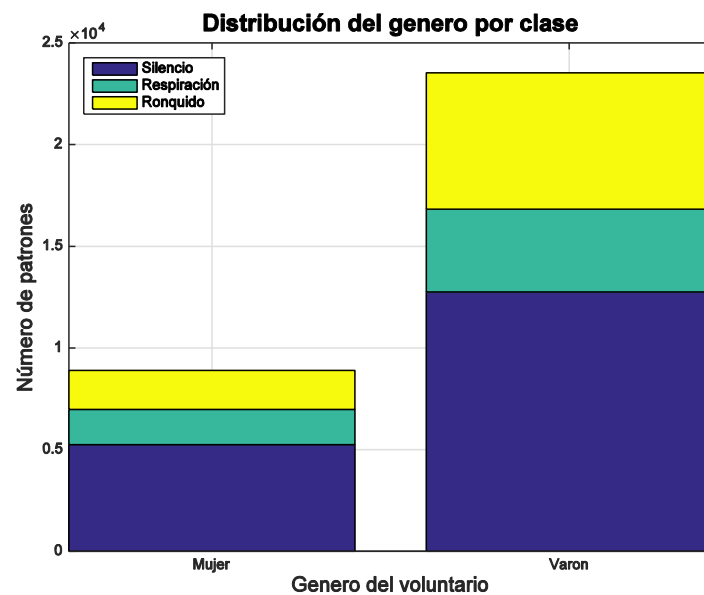


Figura 4-18: Distribución de genero de los voluntarios por clase de evento respiratorio

Para analizar cómo pueden mejorar la clasificación, se eligen los dos clasificadores con mejor puntuación en el paso anterior:

1. Árbol J48
2. Random Forest de 100 árboles

El método consiste en ir añadiendo atributos clínicos a las cuatro características de audio y ver cómo evoluciona la clasificación.

Los resultados se muestran en la Tabla 4-11 para Random Forest y en la Tabla 4-12 para árbol J48.

Atributos comunes: C_k Lp_k Z_k Es_k $Clase_k$

Clasificador: RandomForest 100 árboles. Parámetros: -l 100 -K 0 -S 1

Atributos adicionales	% Clasif. Correctos	True Positive Rate (Sensibilidad)				False Positive Rate (1 – Especificidad)				ROC Área (AUC)			
		Sil.	Res.	Ron.	W.A.	Sil.	Res.	Ron.	W.A.	Sil.	Res.	Ron.	W.A.
Edad	86.7858	0.947	0.66	0.84	0.868	0.142	0.056	0.032	0.098	0.916	0.959	0.962	0.952
Género	85.8879	0.943	0.628	0.837	0.859	0.15	0.056	0.039	0.104	0.955	0.905	0.959	0.947
BMI	86.8674	0.948	0.666	0.839	0.869	0.141	0.054	0.034	0.097	0.958	0.915	0.961	0.951
Peso	86.7676	0.948	0.661	0.837	0.868	0.143	0.054	0.034	0.098	0.958	0.951	0.962	0.952
Altura	86.9281	0.948	0.662	0.842	0.869	0.895	0.73	0.899	0.867	0.959	0.951	0.961	0.952
Peso Altura	86.9581	0.948	0.663	0.844	0.87	0.142	0.052	0.034	0.098	0.959	0.916	0.962	0.952
BMI Edad	86.8765	0.948	0.657	0.844	0.869	0.143	0.052	0.034	0.098	0.959	0.916	0.963	0.953
Peso Edad	86.9309	0.949	0.663	0.841	0.869	0.142	0.053	0.033	0.867	0.959	0.916	0.963	0.952
Altura Edad	86.8402	0.95	0.658	0.838	0.868	0.145	0.051	0.034	0.099	0.959	0.915	0.963	0.952
Peso Género	87.0941	0.949	0.662	0.846	0.871	0.141	0.052	0.033	0.096	0.959	0.916	0.962	0.952
Altura Género	87.0941	0.949	0.662	0.849	0.871	0.141	0.052	0.033	0.096	0.959	0.916	0.962	0.952
Peso Género Edad	87.0034	0.95	0.664	0.841	0.87	0.144	0.052	0.032	0.098	0.96	0.916	0.963	0.953
Peso Altura Edad	87.0851	0.95	0.666	0.842	0.871	0.142	0.052	0.032	0.097	0.96	0.917	0.964	0.953
BMI Género	86.7858	0.947	0.659	0.841	0.868	0.144	0.054	0.033	0.099	0.959	0.915	0.963	0.952
Peso Altura Género	86.183	0.947	0.659	0.842	0.868	0.144	0.053	0.034	0.099	0.959	0.917	0.963	0.953
BMI Género Edad	87.0125	0.948	0.671	0.84	0.87	0.142	0.054	0.031	0.097	0.96	0.916	0.963	0.953
Peso Altura Género Edad	86.94	0.95	0.66	0.84	0.869	0.144	0.052	0.033	0.098	0.96	0.917	0.963	0.953
BMI Peso Altura Género Edad	87.0488	0.948	0.665	0.845	0.87	0.141	0.052	0.033	0.097	0.96	0.917	0.963	0.953

Tabla 4-11: Clasificación de eventos respiratorios con RandomForest añadiendo atributos STOP-BANG

En la Tabla 4-11 correspondiente a RandomForest, se observa que el añadir cualquier atributo del formulario STOP-BANG mejora el porcentaje de clasificación (de 85,26% a ser superior a 86,18%) y la calidad del clasificador (AUC de 0.944 a ser superior a 0.95).

Los conjuntos de atributos que mejor funcionan con el conjunto de voluntarios analizados son Peso/Género, Altura/Género y Peso/Altura/Edad.

Se comprobó cómo mejoraban el rendimiento del árbol J48 al añadir estos tres grupos de atributos. Los resultados se muestran en la Tabla 4-12. Comprobamos que el rendimiento del clasificador J48 mejora respecto a utilizar solo los atributos de audio, pero su rendimiento sigue estando por debajo del Random Forest.

Atributos comunes: C_k Lp_k Z_k Es_k $Clase_k$

Clasificador: árbol J48. Parámetros: -C 0.25 -M 2

Atributos adicionales	% Clasif. Correctos	True Positive Rate (Sensibilidad)				False Positive Rate (1 – Especificidad)				ROC Área (AUC)			
		Sil.	Res.	Ron.	W.A.	Sil.	Res.	Ron.	W.A.	Sil.	Res.	Ron.	W.A.
Peso Género	85.3709	0.94	0.643	0.814	0.854	0.153	0.064	0.036	0.106	0.937	0.89	0.928	0.921
Altura Género	85.4435	0.945	0.627	0.816	0.854	0.157	0.061	0.035	0.108	0.94	0.857	0.929	0.922
Peso Altura Edad	85.5251	0.942	0.647	0.813	0.855	0.15	0.062	0.037	0.105	0.936	0.852	0.932	0.92

Tabla 4-12: Clasificación de eventos respiratorios con J48 añadiendo atributos STOP-BANG

Hasta ahora se ha estudiado cómo clasificar los eventos respiratorios en tres clases: silencio, respiración y ronquido. Si nos interesa centrarnos en la identificación del ronquido (para realizar operaciones de estimulación del paciente, por ejemplo), sería de interés analizar cómo se realiza la clasificación reduciendo el número de clases a dos: no_ronquido vs. ronquido. Para ello, agrupamos las clases silencio y respiración en la clase no_ronquido. Esto se realiza en la práctica fácilmente a través de la tabla '*clase*' de la BBDD (ver anexo 0), en la que ya hay definidos campos para realizar estas agrupaciones. En la extracción de datos, seleccionaremos el campo del grupo que queremos utilizar en lugar del campo de la clase.

Los resultados de esta nueva aproximación quedan reflejados en la Tabla 4-13. Se observa que de manera general todos los clasificadores tienen un desempeño mejor que con tres clases. Los clasificadores con mejor desempeño siguen siendo el árbol J48 y RandomForest de 100 árboles, superando ambos un éxito de clasificación de 91,7% con una AUC mayor de 0,93.

Atributos seleccionados: C_k Lp_k Z_k Es_k $grupo1_k$

Clase Clasif.	Parám. Clasif.	% Clasif. Correctos	True Positive Rate (Sensibilidad)			False Positive Rate (1 – Especificidad)			ROC Área (AUC)		
			NO-RON	RON	W.A.	NO-RON	RON	W.A.	NO-RON	RON	W.A.
ZeroR	n.a.	73.6352	1	0	0.737	1	0	0.737	0.5	0.5	0.5
NaiveBayes Simple	n.a	80.4916	0.945	0.414	0.805	0.586	0.055	0.446	0.831	0.831	0.831
Logistic	-R 1.0E-8 -M -1	84.5184	0.973	0.489	0.845	0.511	0.027	0.383	0.764	0.764	0.764
Multilayer Perceptron 1 capa oculta 3 neuronas	-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H 3	89.7424	0.956	0.733	0.897	0.267	0.044	0.208	0.908	0.908	0.908
Multilayer Perceptron 1 capa oculta 5 neuronas	-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H 5	91.7831	0.978	0.749	0.918	0.251	0.022	0.191	0.93	0.93	0.93
Multilayer Perceptron 1 capa oculta 6 neuronas	-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H 6	91.91	0.971	0.775	0.919	0.225	0.029	0.173	0.944	0.944	0.944
Multilayer Perceptron 1 capa oculta 8 neuronas	-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H 8	92.2547	0.977	0.77	0.923	0.23	0.023	0.176	0.951	0.951	0.951
SMO (SVM) Kernel Lineal	-C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K PolyKernel -C 250007 -E 1.0	84.1194	0.977	0.462	0.841	0.538	0.023	0.403	0.719	0.719	0.719
SMO (SVM) Kernel Lineal Low Order	-C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K PolyKernel -C 250007 -E 1.0 -L	84.1194	0.977	0.462	0.841	0.538	0.023	0.403	0.719	0.719	0.719
J48	-C 0.25 -M 2	91.7105	0.964	0.785	0.917	0.215	0.036	0.168	0.931	0.931	0.931
Random Tree	-K 0 -M 1.0 -S 1	88.9987	0.926	0.789	0.89	0.211	0.074	0.175	0.858	0.858	0.858
Random Forest 100 árboles	-I 100 -K 0 -S 1	92.8351	0.967	0.821	0.928	0.179	0.033	0.141	0.959	0.959	0.959

Tabla 4-13: Resultados de los distintos clasificadores para los eventos ronquido/no_ronquido

Siguiendo la misma aproximación que con dos clases, se analiza cómo influye en la clasificación el incorporar atributos del formulario STOP-BANG. Se analiza con el clasificador con mejor desempeño en la clasificación utilizando solo los atributos de la señal de audio.

Los resultados de esta análisis se recogen en la Tabla 4-14.

Atributos comunes: C_k Lp_k Z_k Es_k grupo1_k

Clasificador: RandomForest 100 árboles. Parámetros: -l 100 -K 0 -S 1

Atributos adicionales	% Clasif. Correctos	True Positive Rate (Sensibilidad)			False Positive Rate (1 – Especificidad)			ROC Área (AUC)		
		NO-RON	RON	W.A.	NO-RON	RON	W.A.	NO-RON	RON	W.A.
Edad	93.47	0.973	0.827	0.935	0.173	0.027	0.134	0.963	0.963	0.963
Género	93.1344	0.969	0.826	0.931	0.174	0.031	0.136	0.961	0.961	0.961
BMI	93.3974	0.973	0.825	0.934	0.175	0.027	0.136	0.962	0.962	0.962
Peso	93.4428	0.973	0.827	0.934	0.173	0.027	0.135	0.962	0.962	0.962
Altura	93.4518	0.972	0.829	0.935	0.171	0.028	0.134	0.962	0.962	0.962
Peso Altura	93.3884	0.972	0.827	0.934	0.173	0.028	0.135	0.962	0.962	0.962
BMI Edad	93.4972	0.974	0.827	0.935	0.173	0.026	0.134	0.963	0.963	0.963
Peso Edad	93.4065	0.973	0.825	0.934	0.175	0.027	0.136	0.962	0.962	0.962
Altura Edad	93.4246	0.973	0.826	0.934	0.174	0.027	0.135	0.963	0.963	0.963
Peso Género	93.3521	0.972	0.827	0.934	0.173	0.028	0.135	0.962	0.962	0.962
Altura Género	93.4065	0.974	0.823	0.934	0.177	0.026	0.133	0.961	0.961	0.961
Peso Género Edad	93.3793	0.972	0.826	0.934	0.174	0.028	0.136	0.963	0.963	0.963
Peso Altura Edad	93.5425	0.974	0.829	0.935	0.171	0.026	0.133	0.962	0.962	0.962
BMI Género	93.4065	0.973	0.824	0.934	0.176	0.027	0.136	0.962	0.962	0.962
Peso Altura Género	93.47	0.973	0.828	0.935	0.172	0.027	0.134	0.963	0.963	0.963
BMI Género Edad	93.4428	0.973	0.826	0.934	0.174	0.027	0.135	0.963	0.963	0.963
Peso Altura Género Edad	93.5153	0.973	0.83	0.935	0.17	0.027	0.132	0.963	0.963	0.963
BMI Peso Altura Género Edad	93.4881	0.974	0.826	0.935	0.174	0.026	0.135	0.963	0.963	0.963

Tabla 4-14: Clasificación de ronquido/no_ronquido con RandomForest más STOP-BANG

Se observa que de nuevo todos los clasificadores presentan mejor desempeño al añadir atributos del formulario STOP-BANG. En este caso, el grupo de atributos que mejor desempeño aporta es el compuesto de Peso/Altura/Edad, con un porcentaje de clasificación correcta de 93,5425%, Especificidad = 0.974, Sensibilidad = 0.867 y AUC = 0.962.

Las diferencias entre incluir BMI o Peso/Altura como atributos de clasificación, puede ser debido a que aunque BMI es una combinación de peso y altura, la relación de BMI-Altura es cuadrática en lugar de lineal (ver ecuación (1) de la sección 3.1.7). En el grupo de voluntarios estudiados, el clasificador tiene mejor desempeño al utilizar Peso/Altura en lugar de BMI.

4.3 Predicción de pacientes con SAHS usando las características de audio

Diversos estudios parten de la hipótesis de que el SAHS modifica las características físicas del tracto vocal, y por lo tanto modificarán las características de la voz del paciente con SAHS frente a un individuo sano (Montero Benavides et al., 2014, 2015). Estas modificaciones podrían ser detectadas usando técnicas de reconocimiento de voz (Goldshtein et al., 2011). En este trabajo partimos de la hipótesis de que los ronquidos se producen de la misma manera que la voz, por lo que se podría detectar a los pacientes que sufren de SAHS mediante el estudio de las características de sus ronquidos (A Roebuck et al., 2014).

Tal y como se detalló en la sección 4.1, en el grupo de estudio hay dos voluntarios con SAHS diagnosticado. Dentro de la BBDD se recoge en la ficha del paciente si está diagnosticado de SAHS, por lo que este campo nos sirve como especificación de la clase. En las siguientes figuras se muestra la distribución de los atributos C_k , Lp_k , Z_k y Es_k frente al diagnóstico de SAHS.

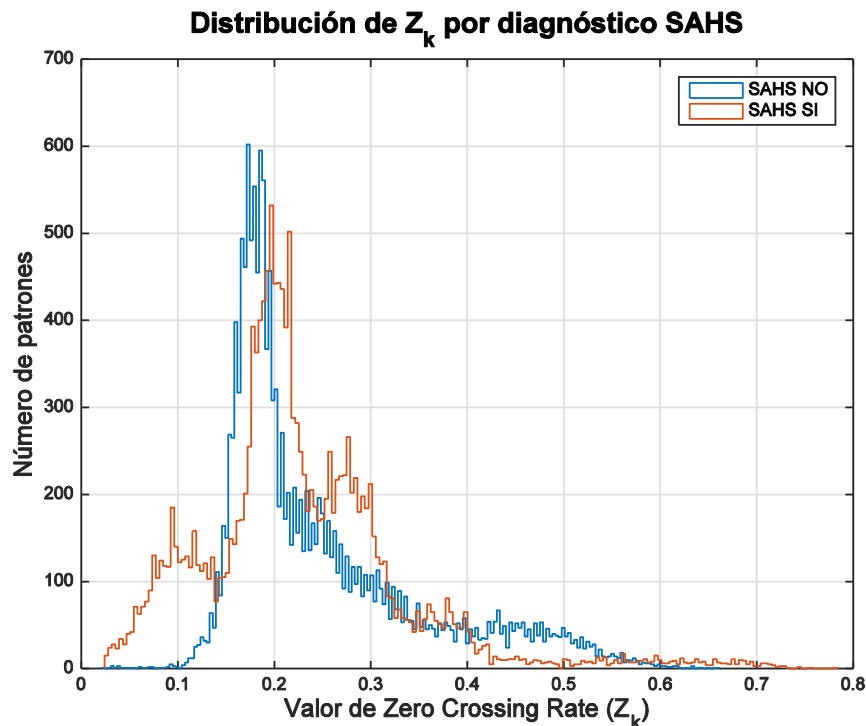


Figura 4-19: Distribución de Z_k por diagnóstico de SAHS

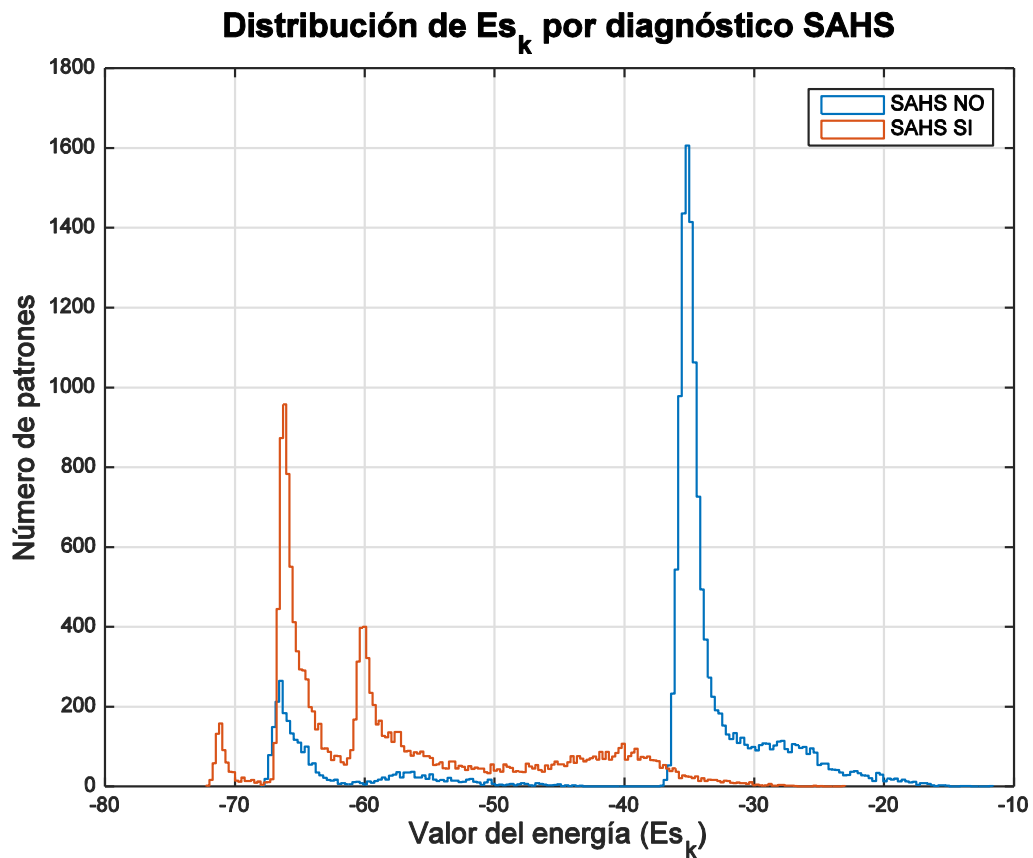


Figura 4-20: Distribución de Es_k por diagnóstico de SAHS

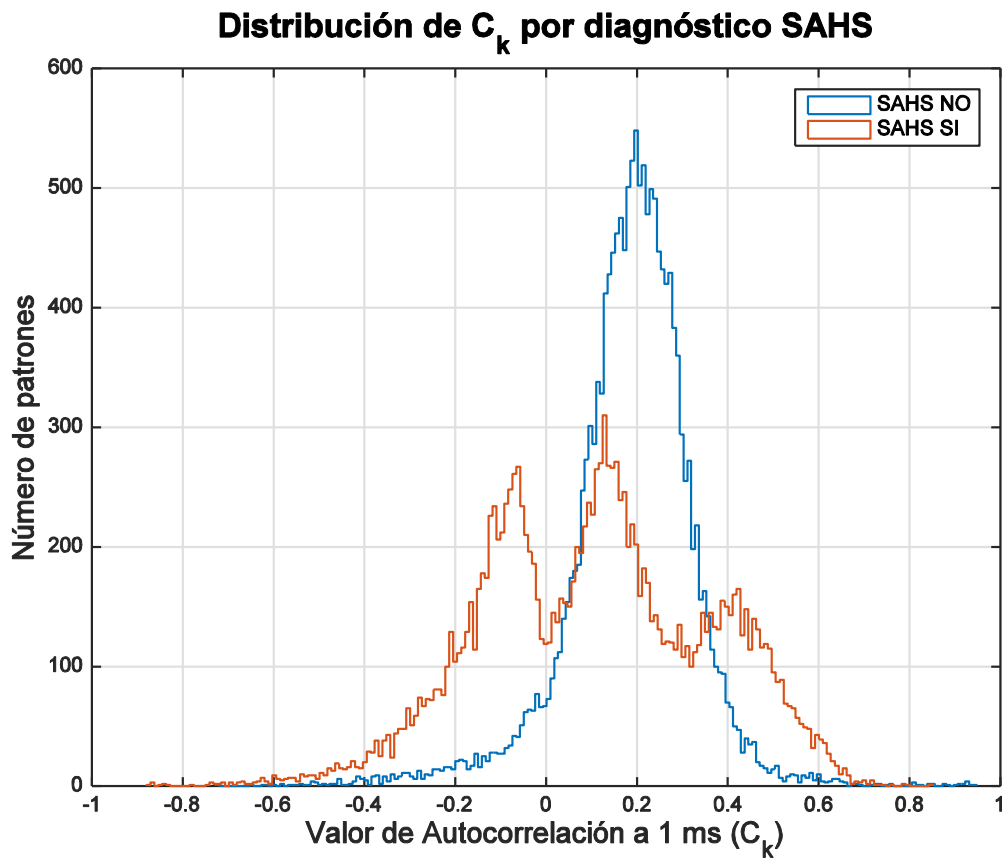


Figura 4-21: Distribución de C_k por diagnóstico de SAHS

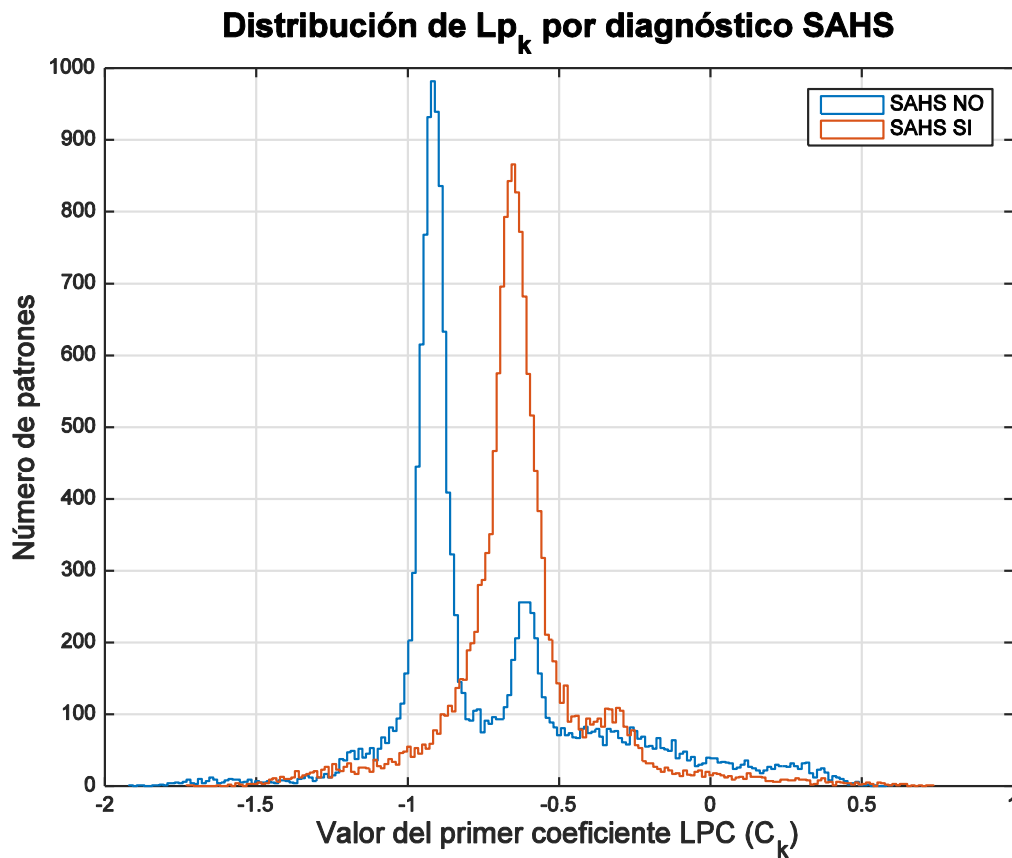


Figura 4-22: Distribución de Lp_k por diagnóstico de SAHS

Se aprecia que las distribuciones presentan separación clara en Es_k (Figura 4-20) y Lp_k (Figura 4-22), y en menor medida en C_k (Figura 4-21) y Z_k (Figura 4-19), por lo que a primera vista parecen buenos candidatos como atributos de clasificación de SAHS. La comparación visual indica que será más fácil la clasificación de SAHS/NO_SAHS.

En la Tabla 4-15 se recogen los resultados de la clasificación con distintos modelos.

Atributos seleccionados: C_k Lp_k Z_k Es_k osa_diag_k

Clase Clasif.	Parám. Clasif.	% Clasif. Correctos	True Positive Rate (Sensibilidad)			False Positive Rate (1 – Especificidad)			ROC Área (AUC)		
			NO-SAHS	SAHS	W.A.	NO-SAHS	SAHS	W.A.	NO-SAHS	SAHS	W.A.
ZeroR	n.a.	50.4263	0	1	0.504	0	4	0.504	0.5	0.5	0.5
NaiveBayes Simple	n.a.	84.201	0.839	0.845	0.842	0.155	0.161	0.158	0.922	0.922	0.922
Logistic	-R 1.0E-8 -M -1	88.1009	0.874	0.888	0.881	0.112	0.126	0.119	0.927	0.927	0.927
Multilayer Perceptron 1 capa oculta 3 neuronas	-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H 3	93.6604	0.94	0.933	0.937	0.067	0.06	0.063	0.987	0.987	0.987
Multilayer Perceptron 1 capa oculta 4 neuronas	-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H 4	95.6013	0.947	0.965	0.956	0.035	0.053	0.044	0.991	0.991	0.991
Multilayer Perceptron 1 capa oculta 8 neuronas	-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H 8	96.3541	0.973	0.955	0.964	0.045	0.027	0.036	0.994	0.994	0.994
SMO (SVM) Kernel Lineal	-C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K PolyKernel -C 250007 -E 1.0	88.9806	0.877	0.903	0.89	0.097	0.123	0.11	0.89	0.89	0.89
J48	-C 0.25 -M 2	96.7985	0.968	0.968	0.968	0.032	0.032	0.32	0.981	0.981	0.981
Ramdom Tree	-K 0 -M 1.0 -S 1	96.209	0.964	0.96	0.962	0.04	0.036	0.038	0.962	0.962	0.962
Ramdom Forest 100 árboles	-I 100 -K 0 -S 1	97.3245	0.973	0.974	0.973	0.026	0.027	0.027	0.996	0.996	0.996

Tabla 4-15: Resultados de los distintos clasificadores para la predicción de SAHS

Se observa que de nuevo que RandomForest es el que mejor desempeño presenta para los parámetros escogidos en los distintos clasificadores utilizados y con el actual grupo de voluntarios. En la tabla Tabla 4-16 se comparan los cuatro clasificadores con mejor puntuación:

	RandomForest	J48	Perceptrón 8 n.	RamdomTree
Sensibilidad	0.973	0.968	0.964	0.962
Especificidad	0.973	0.968	0.964	0.962
AUC	0.996	0.981	0.994	0.962
% Clasif. Correcto	97.3245%	96.7985%	96.3541%	96.209%

Tabla 4-16: Comparación de clasificadores para la predicción de SAHS

Siguiendo con el mismo método que en la sección 4.2, se puede analizar cómo mejora la clasificación al incorporar atributos de los formularios STOP-BANG.

Partiendo de los mismos atributos que se definieron en la sección 4.2, se analiza su distribución en función de la clase de SAHS diagnosticado. En las siguientes figuras se muestran estas distribuciones.

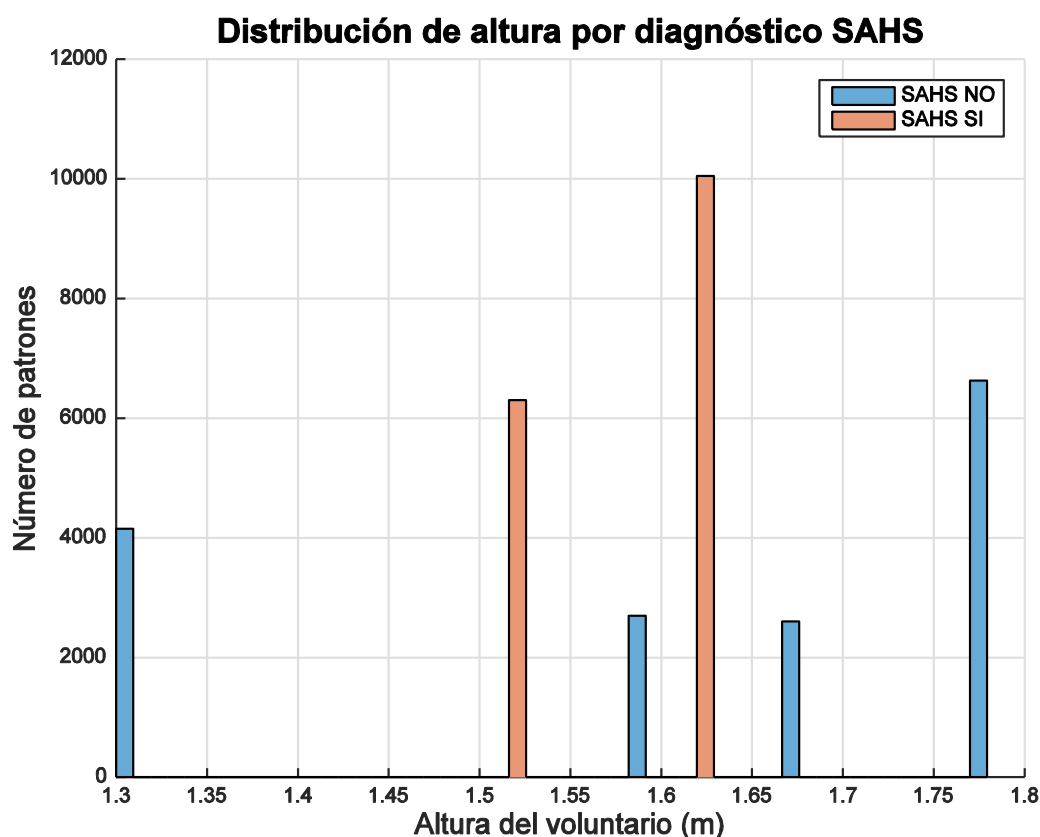


Figura 4-23: Distribución de altura de los voluntarios por diagnóstico de SAHS

Debido al bajo número de voluntarios, la utilización de estos atributos extraídos de los formularios puede quedar distorsionada.

En la distribución del diagnóstico por edad de los voluntarios (Figura 4-25) se observa una clara separación del diagnóstico por edad: los mayores de 40 años presentan diagnóstico de SAHS, los menores de 30 años no presentan diagnóstico de SAHS. Al tener esta frontera tan clara, se va a obtener un 100% de éxito en la clasificación por diagnóstico. Lo mismo sucede con la distribución del BMI (Figura 4-26). Por este motivo, ambos atributos (edad y BMI) se descartan en la clasificación.

Al descartar BMI, es de sospechar que altura (Figura 4-23) y peso (Figura 4-24) también fueren la clasificación al 100%, pero a priori no se descartan.

A primera vista, la distribución del género (Figura 4-27) no parece presentar este problema de frontera clara. Por lo tanto, se continúan las pruebas añadiendo el atributo de género al vector de características de audio.

En la Tabla 4-17 se muestran los resultados de las clasificaciones con este nuevo atributo.

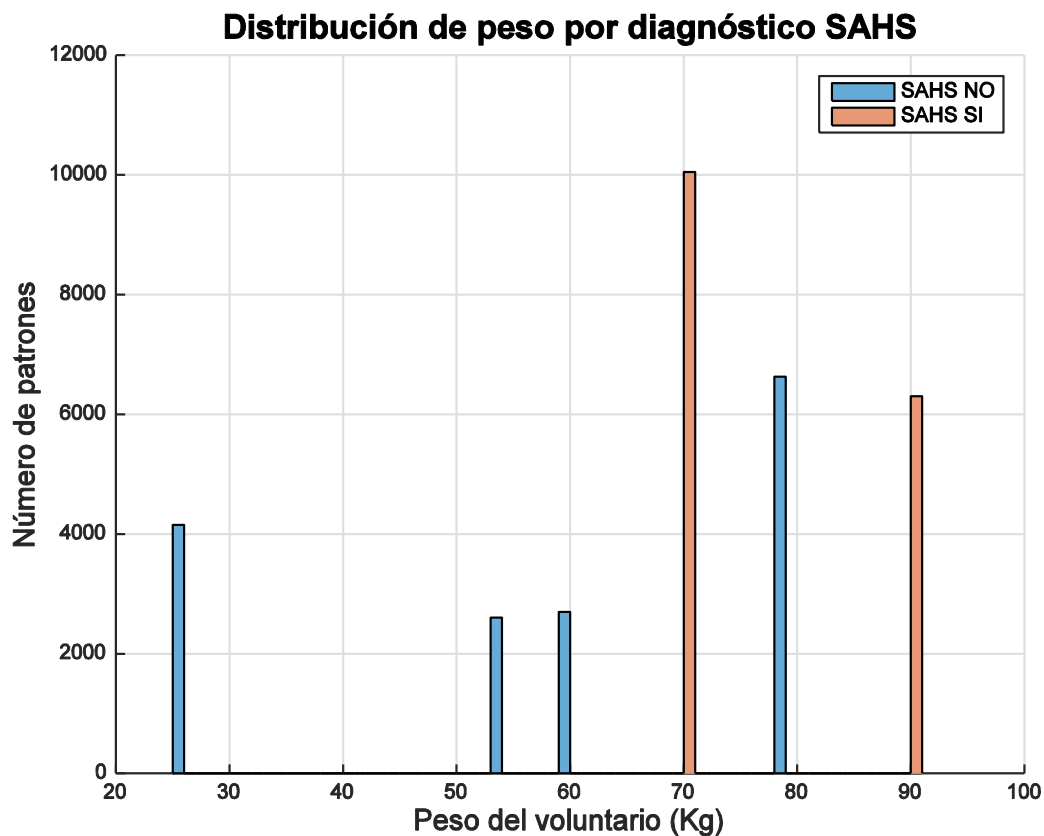


Figura 4-24: Distribución del peso de los voluntarios por diagnóstico de SAHS

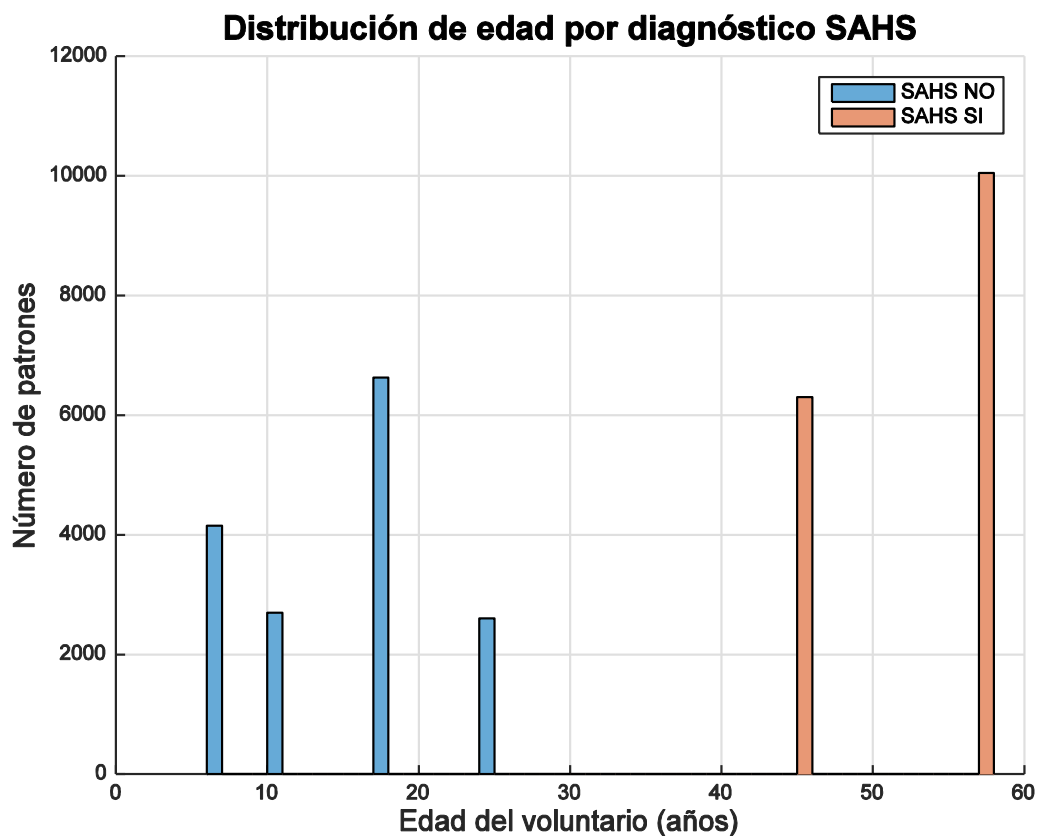


Figura 4-25: Distribución de la edad de los voluntarios por diagnóstico de SAHS

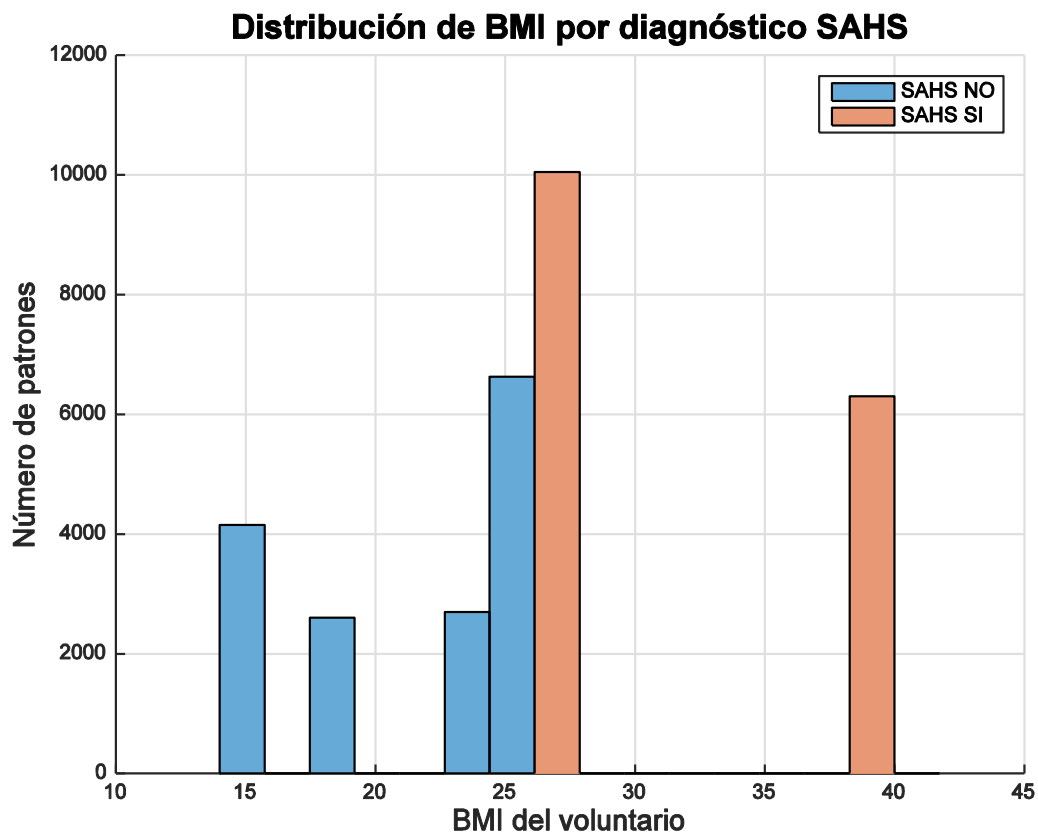


Figura 4-26: Distribución del BMI de los voluntarios por diagnóstico de SAHS

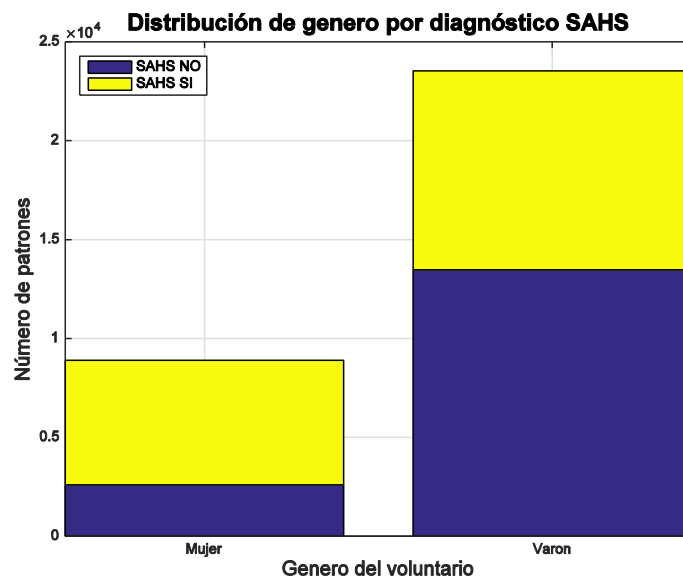


Figura 4-27 Distribución del género de los voluntarios por diagnóstico de SAHS

Atributos seleccionados: C_k Lp_k Z_k Es_k $genero_k$ osa_diag_k

Clase Clasif.	Parám. Clasif.	% Clasif. Correctos	True Positive Rate (Sensibilidad)			False Positive Rate (1 – Especificidad)			ROC Área (AUC)		
			NO-SAHS	SAHS	W.A.	NO-SAHS	SAHS	W.A.	NO-SAHS	SAHS	W.A.
ZeroR	n.a.	50.4263	0	1	0.504	0	1	0.504	0.5	0.5	0.5
NaiveBayes Simple	n.a.	84.6272	0.844	0.848	0.846	0.152	0.156	0.154	0.932	0.932	0.932
Logistic	-R 1.0E-8 -M -1	88.2006	0.874	0.89	0.882	0.11	0.126	0.118	0.949	0.949	0.949
Multilayer Perceptron 1 capa oculta 3 neuronas	-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H 3	95.8915	0.969	0.949	0.959	0.051	0.031	0.041	0.991	0.991	0.991
Multilayer Perceptron 1 capa oculta 4 neuronas	-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H 4	96.9708	0.971	0.969	0.97	0.031	0.029	0.03	0.994	0.994	0.994
Multilayer Perceptron 1 capa oculta 8 neuronas	-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H 8	97.6419	0.974	0.979	0.976	0.021	0.026	0.024	0.997	0.997	0.997
SMO (SVM) Kernel Lineal	-C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K PolyKernel -C 250007 -E 1.0	89.6336	0.884	0.909	0.896	0.091	0.116	0.104	0.896	0.896	0.896
J48	-C 0.25 -M 2	97.5422	0.971	0.98	0.975	0.02	0.029	0.025	0.986	0.986	0.986
Ramdom Tree	-K 0 -M 1.0 -S 1	96.8438	0.965	0.972	0.968	0.028	0.035	0.032	0.968	0.968	0.968
Ramdom Forest 100 árboles	-I 100 -K 0 -S 1	97.9503	0.978	0.981	0.98	0.019	0.022	0.021	0.997	0.997	0.997

Tabla 4-17: Resultados de los distintos clasificadores para la predicción de SAHS incluyendo el género

Al incluir el atributo de género, los modelos de clasificación mejoran, igual que ocurría con la clasificación de eventos respiratorios. El modelo de clasificación que presenta mayor ganancia en la clasificación es el Perceptrón de 1 capa oculta con 8 neuronas, que pasa de un éxito de clasificación del 96.3541% al 97.6419%, pasando su AUC de 0,994 a 0.997.

A continuación, se comprueba el aporte de los atributos de altura y peso. Los resultados se muestran en la Tabla 4-18 y en la Tabla 4-19.

Atributos seleccionados: C_k Lp_k Z_k Es_k altura_k osa_diag_k

Clase Clasif.	Parám. Clasif.	% Clasif. Correctos	True Positive Rate (Sensibilidad)			False Positive Rate (1 – Especificidad)			ROC Área (AUC)		
			NO-SAHS	SAHS	W.A.	NO-SAHS	SAHS	W.A.	NO-SAHS	SAHS	W.A.
Multilayer Perceptron 1 capa oculta 3 neuronas	-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H 3	99.6282	0.993	0.999	0.996	0.001	0.007	0.004	0.998	0.998	0.998
Multilayer Perceptron 1 capa oculta 4 neuronas	-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H 4	99.6282	0.993	0.999	0.996	0.001	0.007	0.004	0.998	0.998	0.998
Multilayer Perceptron 1 capa oculta 5 neuronas	-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H 5	100	1	1	1	0	0	0	1	1	1
J48	-C 0.25 -M 2	99.9909	1	1	1	0	0	0	1	1	1
Ramdom Tree	-K 0 -M 1.0 -S 1	99.9909	1	1	1	0	0	0	1	1	1
Ramdom Forest 100 árboles	-I 100 -K 0 -S 1	100	1	1	1	0	0	0	1	1	1

Tabla 4-18: Resultados de los distintos clasificadores para la predicción de SAHS incluyendo la altura

Al añadir al atributo *altura*, ya se alcanza el 100% de clasificación correcta con un perceptrón de 1 capa oculta y 5 neuronas, en lugar de 8 neuronas.

Atributos seleccionados: C_k Lp_k Z_k Es_k peso_k osa_diag_k

Clase Clasif.	Parám. Clasif.	% Clasif. Correctos	True Positive Rate (Sensibilidad)			False Positive Rate (1 – Especificidad)			ROC Área (AUC)		
			NO-SAHS	SAHS	W.A.	NO-SAHS	SAHS	W.A.	NO-SAHS	SAHS	W.A.
Multilayer Perceptron 1 capa oculta 3 neuronas	-L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H 3	100	1	1	1	0	0	0	1	1	1
J48	-C 0.25 -M 2	99.9547	1	0.999	1	0.001	0	0	1	1	1
Ramdom Tree	-K 0 -M 1.0 -S 1	99.9547	1	0.999	1	0.001	0	0	1	1	1
Ramdom Forest 100 árboles	-I 100 -K 0 -S 1	100	1	1	1	0	0	0	1	1	1

Tabla 4-19: Resultados de los distintos clasificadores para la predicción de SAHS incluyendo el peso

Ambos atributos (altura y peso) mejoran el desempeño de los clasificadores, llegando al 100% tal y como se había sospechado.

En resumen, se puede concluir que las cuatro características de audio seleccionadas (C_k , Lp_k , Z_k y Es_k) son buenos predictores del posible diagnóstico positivo de SAHS, siempre dentro de las limitaciones debidas al bajo número de sujetos de la muestra analizada.

Así mismo, el añadir atributos de los formularios STOP-BANG, mejora la clasificación del diagnóstico. En la Tabla 4-20, se muestra un resumen de los resultados de los clasificadores al añadir el atributo de *género*.

	RandomForest	J48	Perceptrón 8 n.
Sensibilidad	0.981	0.98	0.979
Especificidad	0.979	0.975	0.976
AUC	0.997	0.986	0.997
% Clasif. Correcto	97.9503%	97.5422%	97.6419%

Tabla 4-20: Comparación de clasificadores para la predicción de SAHS usando el *género* del voluntario

En la Tabla 4-21 se muestra la comparación de los resultados de la clasificación del diagnóstico de SAHS sin atributos STOP-BANG y añadiendo el atributo *género* del voluntario.

	RandomForest		J48		Perceptrón 8 n.	
	SIN	CON	SIN	CON	SIN	CON
Sensibilidad	0.973	0.981	0.968	0.98	0.964	0.979
Especificidad	0.973	0.979	0.968	0.975	0.964	0.976
AUC	0.996	0.997	0.981	0.986	0.994	0.997
% Clasif. Correcto	97.3245%	97.9503%	96.7985%	97.5422%	96.3541%	97.6419%

Tabla 4-21: Comparación de los clasificadores en la predicción de SAHS sin y con atributo *género* del voluntario

El clasificador que mayor mejoría presenta es el perceptrón de 1 capa oculta y 8 neuronas. Hay que observar, de todas formas, que los modelos de clasificación sin añadir el atributo *género* ya presentaban un desempeño muy alto.

5 Conclusiones

El SAHS es un trastorno del sueño bastante común, con una prevalencia del 2% al 7% de la población. Se estima que el 90% de los enfermos no están diagnosticados, ya que la prueba diagnóstica de PSG es cara y limitada por el número de camas disponibles en la Unidad del Sueño de los Hospitales. El SAHS presenta un alto impacto social y económico debido a la falta de un descanso adecuado durante el sueño, por lo produce estados de somnolencia durante la vigilia que pueden provocar descenso en la productividad en el trabajo y accidentes en el manejo de maquinaria y vehículos. Así mismo, la situación de hipoxia durante el sueño puede inducir diversas enfermedades circulatorias, neurocognitivas y metabólicas en función de su severidad. Por lo tanto, el poder disponer de un sistema de recomendación de diagnóstico de SAHS de bajo coste, sencillo y que alcance a un gran porcentaje de la población, sería de gran utilidad para los Servicios Médicos implicados en el diagnóstico y tratamiento del SAHS.

La hipótesis inicial de trabajo de este TFM se basa en investigaciones que apuntan a que en los pacientes con SAHS se modifica la forma en que hablan debido a la inflamación de las zonas blandas de la vía aérea superior que se colapsan (Montero Benavides et al., 2014; Solé-Casals et al., 2014). Por lo tanto, es de suponer que si el ronquido se genera en las mismas estructuras que la voz, las personas que sufran de SAHS ronquen de manera distinta que las personas sanas.

En este Trabajo de Fin de Master se ha propuesto un método de análisis de los sonidos respiratorios durante el sueño (silencio – respiración – ronquido), registrados fuera de un entorno controlado de laboratorio con el fin de emitir una recomendación de diagnóstico de SAHS. Estos métodos de análisis son lo suficientemente sencillos como para incorporarlos en una aplicación de *Smartphone* que pueda ser distribuida como herramienta de cribado entre la población, de manera que analice los sonidos respiratorios nocturnos del usuario de la aplicación y emita una recomendación de visitar al médico si se calcula una alta probabilidad de sufrir SAHS. Las características de audio extraídas de los registros nocturnos permiten una clasificación correcta de SAHS/NO-SAHS del 97.33% utilizando el clasificador RandomForest, con una sensibilidad de 0.973, una especificidad de 0.973 y una AUC de 0.996.

Así mismo, las características de audio registradas permiten una clasificación *on-line* de los eventos respiratorios, ya sean Respiración/Silencio/Ronquido o No-ronquido/Ronquido. Esto abre la puerta a desarrollar sistemas que estimulen al usuario cuando ronca para intentar evitar que se produzca este hecho. Algunos tipos de apneas son posicionales (se ronca en posición decúbito supino), por lo que se podría estimular al paciente cuando ronca para que realizara un cambio de postura a una lateral o prono (Oksenberg & Silverberg, 1998). Actualmente se realiza con métodos físicos, como por ejemplo, colocando pelotas de tenis cosidas a lo largo de la espalda en una camiseta y poniéndosela para dormir, pero es rechazada por un elevado número de pacientes por ser incómoda (Dan Anker Hofsoy et al., 2009). En las apneas posicionales se podría llegar a eliminar totalmente el problema manteniendo la postura adecuada (LEE et al., 2009).

Adicionalmente, la incorporación de datos biométricos de los voluntarios recogidos mediante formularios STOP-BANG al modelo de clasificación, mejoran los resultados de

este, tanto en la clasificación de eventos respiratorios como en la recomendación de diagnóstico de SAHS. Debido al número bajo de voluntarios (7), no se pudieron evaluar todos los datos biométricos recogidos, por lo que habrá que seguir estudiando su aporte en futuros trabajos, cuando se disponga de una población de experimentación mayor.

Algunos líneas de investigación de detección de SAHS mediante modificaciones en la voz presentan problemas al aumentar la población de estudio (Montero Benavides et al., 2015), por lo que se sospecha que será necesario realizar futuros trabajos de verificación de los métodos presentados en este TFM para evaluar su rendimiento en muestras amplias de población que representen un escenario clínico real, ya que la población utilizada en este TFM ha sido pequeña (7 voluntarios).

Otra aplicación interesante de los resultados de este TFM es en la PSG pediátrica. Hoy en día se da gran importancia al SAHS en niños (Vlahandonis et al., 2014; Vlahandonis, Nixon, Davey, Walter, & Horne, 2013), aunque su causa es distinta que en adultos, siendo principalmente por hipertrofia adenotonsilar. Hay una gran preocupación tanto sobre los pacientes pediátricos con SAHS como en roncadores primarios, ya que ambos problemas afectan a su desarrollo y conducta (Biggs, Nixon, & Horne, 2014). La realización de una PSG a un niño implica un método bastante incómodo para este y en un entorno bastante agresivo (ver Figura 2-2). El poder utilizar la futura herramienta de cribado basada en *Smartphone* podría indicar a los padres la necesidad de visitar al médico por un problema en su hijo, y al pediatra el disponer de una herramienta para poder realizar un diagnóstico sin necesidad de recurrir a una PSG. En este TFM ya se han utilizado voluntarios con edades de 6 años, pero en futuros trabajos habrá que analizar en mayor profundidad cómo influye la edad en las características de audio y en la clasificación de eventos y diagnósticas de SAHS al disponer de una población mayor.

Para alcanzar los resultados de este TFM, se definieron una serie de objetivos al inicio de este proyecto de Trabajo Fin de Master. Se puede concluir que todos los objetivos han sido alcanzados. A continuación se revisa la consecución de cada objetivo:

1. *Encontrar y definir las características de audio que mejor definan la separación de los tres eventos respiratorios Ronquido – Respiración – Silencio*

Se han definido cuatro atributos de audio (sección 3.2.6), que permiten una clasificación aceptable de estos eventos (sección 4.2), siendo el clasificador con mejor desempeño Random-Forest de todos los modelos probados. Se obtuvo una Tasa de Clasificación Correcta del 85.26% con una sensibilidad de 0.853, especificidad de 0.892 y AUC de 0.944. Indicar que el etiquetado de los eventos para el entrenamiento de los modelos ha sido manual y sobre una muestra de siete voluntarios.

También se ha verificado la clasificación de los eventos en ronquido/no-ronquido (sección 4.2). En este caso, también el clasificador Random-Forest es el que ha obtenido mejor desempeño. Se obtuvo una Tasa de Clasificación Correcta del 92.84% con una sensibilidad de 0.928, una especificidad de 0.859 y AUC de 0.959. Esto permite avanzar hacia el desarrollo de un dispositivo que interactúe con el paciente cuando ronca, que le fuerce a cambiar de postura para evitar el ronquido y colapso de la vía aérea.

En base a los resultados obtenidos, se puede concluir que los atributos definidos y los métodos de clasificación utilizados podrían aplicarse en la práctica clínica diaria. Los test

con una AUC mayor o igual a 0.85 ya se utilizan normalmente en la práctica clínica (Greulich et al., 2013; Reiter et al., 2012).

2. *Realizar una predicción de si el paciente que se está registrando es susceptible de sufrir SAHS, para emitir una recomendación de visita al especialista médico*

El uso de los cuatro atributos de audio definidos para clasificar los voluntarios en SAHS/NO-SAHS, utilizando un modelo de Random-Forest (sección 4.3) produce una Tasa de Clasificación Correcta del 97.33% con una sensibilidad de 0.973, una especificidad de 0.973 y una AUC de 0.996. Estos excelentes resultados animan a pensar en el desarrollo de una aplicación de teléfono móvil que realice una recomendación de visitar al médico en base a la clasificación de SAHS/NO-SAHS del paciente.

El modelo de clasificación SAHS/NO-SAHS es susceptible de realizarse *off-line* debido a su metodología, por lo que los requerimientos de capacidad de computación son menores que en el caso de detección de eventos respiratorios.

3. *Seleccionar los datos biométricos que ayuden a mejora la clasificación de las características de audio y la predicción de la posibilidad de sufrir SAHS*

La utilización de datos de los voluntarios recogidos en los formularios STOP-BANG mejoran los modelos de clasificación utilizados basados en los atributos de las ventanas de audio, tanto para la clasificación de eventos respiratorios como el diagnóstico de SAHS (secciones 4.2 y 4.3).

Debido al bajo número de voluntarios, no se ha podido evaluar la aportación de todos los tipos de datos recogidos, ya que algunos presentaban distribuciones de valores no válidas para su utilización.

4. *Consolidar toda la información adquirida, datos y resultados, en una base de datos que conforme los cimientos del sistema de recogida de datos masivo y on-line*

Se ha diseñado e implementado un primer esquema de base de datos (ver anexo A) sobre PostgreSQL 9.4, y rutinas en Matlab® para la inserción y extracción de datos (anexo A y secciones 3.1.6, 3.1.7, 3.1.8, 3.1.9, 3.2.7 y 3.3.1).

Esta base de datos forma el núcleo inicial sobre el que desarrollar la aplicación de recogida de datos masiva y *on-line* basada en *Smartphone*. También permite el acceso concurrente a los datos por distintos investigadores, así como gestionar los permisos y trazas de auditoría de acceso para cumplir con la normativa vigente.

Su estructura permite añadir fácilmente nuevos campos de atributos sobre los segmentos y ventanas de audio ya etiquetadas, así como añadir nuevas etiquetas de eventos.

5. *Contactar con los equipos médicos relacionados con el diagnóstico y tratamiento de SAHS*

Durante la realización de este Trabajo Fin de Máster se han comenzado contactos para colaborar con distintos hospitales en desarrollo de nuevas líneas de trabajo basadas en este

estudio, de manera que se pueda acceder a los datos clínicos y de diagnóstico de sus pacientes de SAHS:

- Hospital Fundación de Alcorcón:
 - Departamento de Otorrinolaringología
 - Unidad del Sueño
 - Departamento de Neumología
- Se han iniciado contactos con otros tres hospitales

En los registros de los voluntarios no se han encontrado marcados periodos de apnea. Este aspecto se podrá estudiar en cuanto se disponga de los registros clínicos.

Hay que reseñar que los resultados obtenidos en este Trabajo Fin de Master son a partir de una muestra muy pequeña (siete voluntarios, 28 sesiones de registro), por lo que habrá que estudiar cómo evolucionan los modelos de clasificación y las características utilizadas con una población más amplia y variada, que refleje un escenario clínico real.

6 Trabajo futuro

Partiendo de que el presente Trabajo Fin de Master constituye el punto de inicio para la realización de la Tesis Doctoral del autor de este trabajo, se pueden definir ya algunas líneas de trabajos futuros para la mejora de los métodos y técnicas utilizados, así como nuevas investigaciones que ayuden a la realización de la Tesis.

Tal y como se apuntó a lo largo de este documento, quedan algunos puntos abiertos en la metodología aplicada:

- Analizar la posible aplicación de filtros de reducción de ruido
- Optimización de los clasificadores utilizados
- Análisis de la correlación Lp_k / Z_k , y ver si es factible eliminar el atributo con mayor coste computacional

Un mejora importante consistiría en automatizar lo máximo posible la asignación de etiquetas de evento a las ventanas de audio. Esta es a día de hoy una operación altamente manual, con el consiguiente riesgo de producir errores. Se plantea el desarrollar una aplicación para estación de trabajo que permita simplificar este proceso mediante:

- Conexión directa a la BBDD para la gestión de todas las tablas
- Gestión de los datos de pacientes, dispositivos de grabación
- Definición de los datos de los registros de audio y carga de segmentos estos
- Definición de los segmentos $s(n)$ a analizar y su carga en la BBDD
- Mostrar la forma de onda y el espectro del segmento $s(n)$ que se esté analizando
- En base a los clasificadores ya entrenados en este TFM, sugerir una etiqueta de clasificación para cada ventana $s_k(n)$
- Registro de la etiqueta de clasificación definitiva asociada a cada ventana $s_k(n)$
- Cálculo de las características de las ventanas $s_k(n)$ y su carga en la BBDD
- Diseño modular para poder añadir el cálculo de nuevas características, filtros, etc.

Esto nos permitiría procesar un mayor volumen de registros de audio a obtener de la colaboración con los hospitales, con el objetivo de mejorar la eficacia del diagnóstico incrementando al máximo el número y diversidad de los patrones utilizados en el entrenamiento de los clasificadores. Por lo tanto, también es necesario aumentar el número de hospitales implicados en el proyecto.

Para la obtención de datos a gran escala, es necesario desarrollar una aplicación de *Smartphone* que registre los sonidos de la respiración durante el sueño, realice los cálculos y los envíe a un servidor central para su incorporación en la BBDD. Todo esto conlleva la implantación de la infraestructura necesaria para poder gestionar las conexiones de los *Smartphones* (servidores de aplicaciones, firewalls, certificados, etc.) así como ubicar la BBDD en un servidor con la potencia y capacidad de almacenamiento suficiente para absorber este crecimiento.

Hasta ahora solo se ha centrado la recomendación de diagnóstico en la utilización de los sonidos respiratorios durante el sueño. Tal y como se indicó en este trabajo, diversos estudios parten de la hipótesis de que el SAHS modifica las características físicas del tracto vocal, y

por lo tanto modificarán las características de la voz del paciente con SAHS frente a un individuo sano. Estas modificaciones podrían ser detectadas usando técnicas de reconocimiento de voz. Por lo tanto, una posible vía de trabajo sería combinar el análisis de los sonidos respiratorios durante el sueño con el análisis de las modificaciones en la voz para intentar mejorar el diagnóstico de SAHS.

Ya que la incorporación de datos recogidos en los formularios STOP-BANG mejoraron el rendimiento de los clasificadores, otra vía de investigación consistiría en añadir nuevos indicadores de SAHS utilizados en la práctica clínica. Uno de estos indicadores es la escala Mallampati/Friedman, basada en medidas antropométricas de la cavidad bucal.

La escala Mallampati fue inicialmente descrita en anestesiología para predecir la facilidad de intubación de los pacientes (Mallampati et al., 1985). Esta clasificación se basa en la anatomía observada de la cavidad bucal, especialmente en la visibilidad de la base de la úvula, istmo de las fauces y el paladar blando, con la lengua protruida. Las clases definidas se pueden observar en la Figura 6-1.

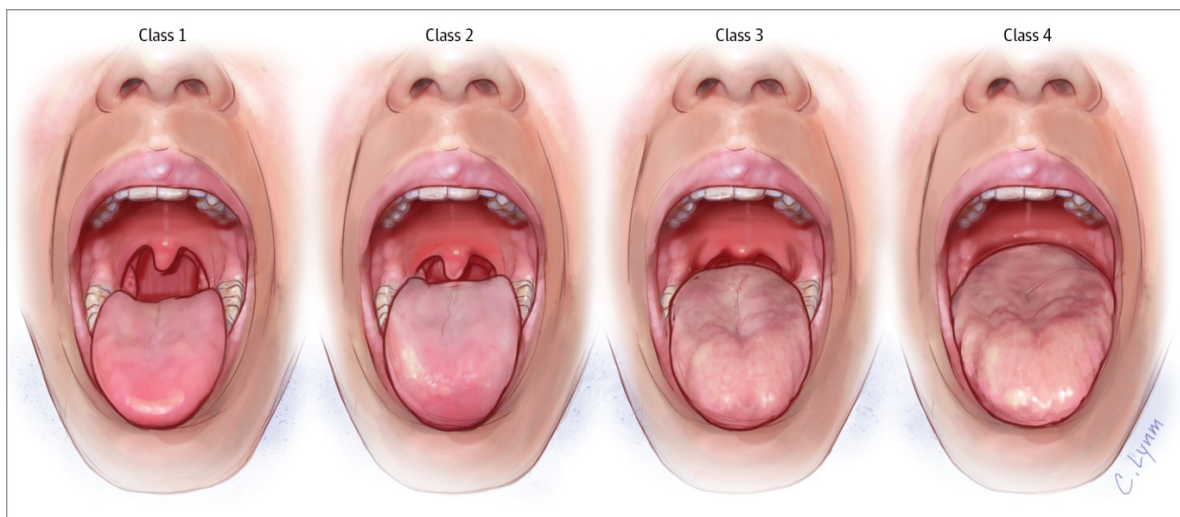


Figura 6-1: Clasificación de Mallampati¹²

Posteriormente, Friedman asocia esta clasificación con la predicción de SAHS en base a modificaciones en el índice Mallampati original, observando la cavidad oral con la lengua sin protruir (M Friedman et al., 1999). Podemos observar esta nueva clasificación en la Figura 6-2.

Existe una correlación entre la posición de la lengua de Friedman y la posibilidad de sufrir SAHS (Michael Friedman, Hamilton, Samuelson, Lundgren, & Pott, 2013).

¹² Fuente: <http://jama.jamanetwork.com>

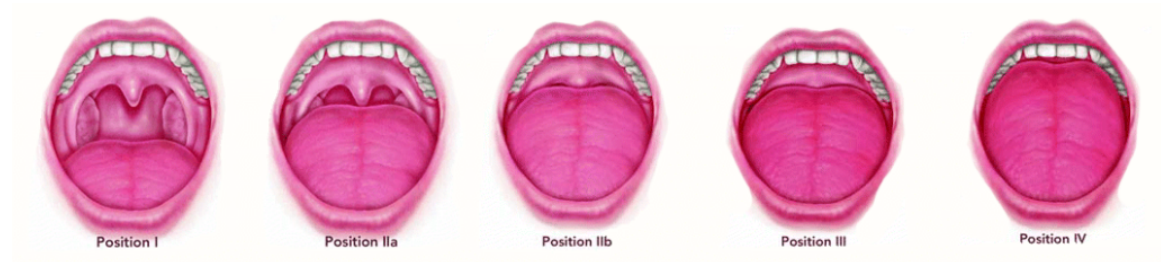


Figura 6-2: Clasificación de Friedman¹³

Una última línea de trabajo sería el desarrollo de un dispositivo portátil para ubicar en la cabeza o cuello del paciente, que permita la recogida de datos y disponga de estimuladores que actúen frente a determinados eventos. Este proyecto enlaza con la detección *on-line* de los eventos respiratorios, con el fin de estimular al paciente para que realice un cambio de postura cuando ronca. Se está trabajando un prototipo de este sistema basado en módulos Arduino. Con este sistema se podrá recoger:

- Posición del paciente
- Sonido traqueal o próximo a la tráquea
- Actividad nocturna del paciente

La conexión con el *Smartphone* se realizaría mediante Bluetooth 4.0 Low Energy. El *Smartphone* analizaría el audio y cuando detecte ronquido puede habilitar los estimuladores para producir un cambio de postura que evite el ronquido. La posición del cuerpo se recogería mediante un acelerómetro de tres ejes.

¹³ Fuente: <http://curemysleepapnea.com>

Referencias

- Abeyratne, U. R., Wakwella, A. S., & Hukins, C. (2005). Pitch jump probability measures for the analysis of snoring sounds in apnea. *Physiological Measurement*, 26(5), 779–798. doi:10.1088/0967-3334/26/5/016
- Abrishami, A., Khajehdehi, A., & Chung, F. (2010). A systematic review of screening questionnaires for obstructive sleep apnea. *Canadian Journal of Anaesthesia = Journal Canadien D'anesthésie*, 57(5), 423–38. doi:10.1007/s12630-010-9280-x
- American Academy of Sleep Medicine. (2001). *International classification of sleep disorders (ICSD), revised: Diagnostic and coding manual*. Chicago, Illinois: American Academy of Sleep Medicine.
- American Association of Sleep Technologist. (2012). Standard Polysomnography. Retrieved from www.aastweb.org
- Bachu, R. G., Kopparthi, S., Adapa, B., & Barkana, B. D. (2010). Separation of Voiced and Unvoiced using Zero crossing rate and Energy of the Speech Signal. In K. Elleith (Ed.), *Advanced Techniques in Computing Sciences and Software Engineering* (pp. 279–282). Springer Netherlands. doi:10.1007/978-90-481-3660-5_47
- Behar, J., Roebuck, A., Domingos, J. S., Geder, E., & Clifford, G. D. (2013). A review of current sleep screening applications for smartphones. *Physiological Measurement*, 34(7), R29–46. doi:10.1088/0967-3334/34/7/R29
- Behar, J., Roebuck, A., Shahid, M., Daly, J., Hallack, A., Palmius, N., ... Clifford, G. D. (2013). SleepAp: an automated obstructive sleep apnoea screening application for smartphones. *Computing in Cardiology 2013 (CinC)*, IEEE Eng. Med. Biol. Soc.
- Benesty, J., Sondhi, M. M., & Huang, Y. (2008). *Springer Handbook of Speech Processing*. (J. Benesty, M. M. Sondhi, & Y. Huang, Eds.). Berlin, Heidelberg: Springer-verlag.
- Benoist, L. B. L., Morong, S., van Maanen, J. P., Hilgevoord, A. A. J., & de Vries, N. (2014). Evaluation of position dependency in non-apneic snorers. *European Archives of Oto-Rhino-Laryngology: Official Journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS): Affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery*, 271(1), 189–94. doi:10.1007/s00405-013-2570-5
- Biggs, S. N., Nixon, G. M., & Horne, R. S. C. (2014). The conundrum of primary snoring in children: What are we missing in regards to cognitive and behavioural morbidity? *Sleep Medicine Reviews*, 18(6), 463–475. doi:10.1016/j.smr.2014.06.009
- Chung, F., Liao, P., Sasaki, E., Elsaid, H., & Yegneswaran, B. (2011). A Higher Score On STOP-Bang Questionnaire Predicts More Severe OSA. *American Journal of Respiratory and Critical Care Medicine*, 183. doi:10.1164/ajrccm-conference.2011.183.1_MeetingAbstracts.A223710.1164/ajrccm-conference.2011.183.1_MeetingAbstracts.A2237
- Chung, F., Subramanyam, R., Liao, P., Sasaki, E., Shapiro, C., & Sun, Y. (2012). High STOP-Bang score indicates a high probability of obstructive sleep apnoea. *British Journal of Anaesthesia*, 108(5), 768–775. doi:10.1093/bja/aes022
- Chung, F., Yang, Y., & Liao, P. (2013). Predictive Performance of the STOP-Bang Score for Identifying Obstructive Sleep Apnea in Obese Patients. *Obesity Surgery*, 23(12), 2050–2057. doi:10.1007/s11695-013-1006-z

- Chung, F., Yegneswaran, B., Liao, P., Chung, S. A., Vairavanathan, S., Islam, S., ... Shapiro, C. M. (2008). STOP questionnaire: a tool to screen patients for obstructive sleep apnea. *Anesthesiology*, 108(5), 812–21. doi:10.1097/ALN.0b013e31816d83e4
- Counter, P., & Wilson, J. A. (2004). The management of simple snoring. *Sleep Medicine Reviews*, 8(6), 433–441. doi:10.1016/j.smr.2004.03.007
- Dafna, E., Tarasiuk, A., & Zigel, Y. (2013). Automatic detection of whole night snoring events using non-contact microphone. *PloS One*, 8(12), e84139. doi:10.1371/journal.pone.0084139
- Dirección General de Tráfico. (2014). *Plan de Investigación en Seguridad Vial y Movilidad 2013-2016*. Madrid - España. Retrieved from <http://www.dgt.es/es/seguridad-vial/investigacion/planes-investigacion/>
- Fiz Fernández, J. A., Solà Soler, J., & Jané Campos, R. (2011). Snoring analysis methods. *Medicina Clínica*, 137(1), 36–42. doi:10.1016/j.medcli.2010.04.004
- Fiz, J. A., Abad, J., Jane, R., Riera, M., Mananas, M. A., Caminal, P., ... Morera, J. (1996). Acoustic analysis of snoring sound in patients with simple snoring and obstructive sleep apnoea. *European Respiratory Journal*, 9(11), 2365–2370. doi:10.1183/09031936.96.09112365
- Friedman, M., Hamilton, C., Samuelson, C. G., Lundgren, M. E., & Pott, T. (2013). Diagnostic value of the Friedman tongue position and Mallampati classification for obstructive sleep apnea: a meta-analysis. *Otolaryngology--Head and Neck Surgery : Official Journal of American Academy of Otolaryngology-Head and Neck Surgery*, 148(4), 540–7. doi:10.1177/0194599812473413
- Friedman, M., Tanyeri, H., La Rosa, M., Landsberg, R., Vaidyanathan, K., Pieri, S., & Caldarelli, D. (1999). Clinical predictors of obstructive sleep apnea. *The Laryngoscope*, 109(12), 1901–1907. doi:10.1097/00005537-199912000-00002
- Furukawa, T., Nakano, H., Hiriyama, K., Tanahashi, T., Yosohara, K., Sudo, N., ... Nishima, S. (2010). Relationship between snoring sound intensity and daytime blood pressure. *Sleep and Biological Rhythms*, 8(4), 245–253. doi:10.1111/j.1479-8425.2010.00455.x
- Goldshtein, E., Tarasiuk, A., & Zigel, Y. (2011). Automatic Detection of Obstructive Sleep Apnea Using Speech Signals. *IEEE Transactions on Biomedical Engineering*, 58(5), 1373–1382. doi:10.1109/TBME.2010.2100096
- Greulich, T., Hattesohl, A., Grabisch, A., Koepke, J., Schmid, S., Noeske, S., ... Koczulla, A. R. (2013). Detection of obstructive sleep apnoea by an electronic nose. *European Respiratory Journal*, 42(1), 145–155. doi:10.1183/09031936.00091712
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software. *ACM SIGKDD Explorations Newsletter*, 11(1), 10. doi:10.1145/1656274.1656278
- Hofsoy, D. A., Clauss, J. F., & Wolf, B. (2009). Monitoring and therapy of sleep-related breathing disorders. In *Wearable Micro and Nano Technologies for Personalized Health (pHealth), 2009* 6th International Workshop on (pp. 41–44).
- Hofsoy, D. A., Clauss, J., & Wolf, B. (2009). An intelligent implant system for monitoring and biofeedback therapy of snoring. In O. Dissel & W. C. Schlegel (Eds.), *MICRO-AND NANOSYSTEMS IN MEDICINE, ACTIVE IMPLANTS, BIOSENSORS* (Vol. 25, pp. 196–199). 233 SPRING STREET, NEW YORK, NY 10013, UNITED STATES: SPRINGER.
- Huan, X., Acero, A., & Hon, H.-W. (2001). *Spoken Language Processing. A guide to theory, algorithm, and system development*. New Jersey: Prentice-Hall, Inc.
- Idzikowski, C. (2012). *Sleep Disorders* (1st ed.). Janeza Trdine 9, 51000 Rijeka, Croatia Copyright: inTech. Retrieved from www.intechopen.com

- Jefatura del Estado. (1999). Ley Orgánica 15 / 1999, de 13 de diciembre, de Protección de Datos de Carácter Personal . TEXTO CONSOLIDADO. *Boletín Oficial Del Estado*, 298(14 de diciembre de 1999), 1–21. Retrieved from http://www.agpd.es/portaIwebAGPD/canaIdocumentacion/legislacion/estatal/common/pdfs/2014/Ley_Organica_15-1999_de_13_de_diciembre_de_Proteccion_de_Datos_Consolidado.pdf
- Judd, S. J. (2010). *Sleep Disorders Sourcebook*. (S. J. Judd, Ed.) (3rd ed., Vol. 2). Omnigraphics, Inc. Retrieved from www.healthreferenceseries.com
- Jurado Gámez, B., Gómez-Chaparro Moreno, J. L., Muñoz Calero, M., Ruiz Laguna, J., Muñoz Cabrera, L., Cosano Povedano, A., & López-Barea, J. (2010). Variation in protein expression depending on the severity of sleep apnoea-hypopnoea syndrome. *Archivos de Bronconeumología*, 46(6), 288–93. doi:10.1016/j.arbres.2009.12.007
- Jurado-Gámez, B., Guglielmi, O., Gude, F., & Buena-Casal, G. (2015). Accidentes laborales, absentismo y productividad en pacientes con apneas del sueño. *Archivos de Bronconeumología*, 51(5), 213–218. doi:10.1016/j.arbres.2014.07.003
- Justicia, M. De, & Ministerio De Justicia. (2012). Real Decreto 1720 / 2007, de 21 de diciembre, por el que se aprueba el Reglamento de desarrollo de la Ley Orgánica 15 / 1999, de 13 de diciembre, de protección de datos de carácter personal . TEXTO CONSOLIDADO. *Boletín Oficial Del Estado*, 17(19 de enero de 2008), 1–53. Retrieved from http://www.agpd.es/portaIwebAGPD/canaIdocumentacion/legislacion/estatal/common/pdfs/2014/Real_Decreto_1720-2007_de_21_de_diciembre_por_el_que_se_aprueba_Reglamento_de_desarrollo_Ley_Organica_15-1999_Consolidado.pdf
- Kang, K., Park, K. S., Kim, J. E., Kim, S. W., Kim, Y. T., Kim, J. S., & Lee, H. W. (2013). Usefulness of the Berlin Questionnaire to identify patients at high risk for obstructive sleep apnea: A population-based door-to-door study. *Sleep and Breathing*, 17(2), 803–810. doi:10.1007/s11325-012-0767-2
- Karunajeewa, A. S., Abeyratne, U. R., & Hukins, C. (2008). Silence-breathing-snore classification from snore-related sounds. *Physiological Measurement*, 29(2), 227–243. doi:10.1088/0967-3334/29/2/006
- Kasai, T., & Bradley, T. D. (2011). Obstructive sleep apnea and heart failure: pathophysiologic and therapeutic implications. *Journal of the American College of Cardiology*, 57(2), 119–27. doi:10.1016/j.jacc.2010.08.627
- Kreivi, H.-R., Salmi, T., Maasilta, P., & Bachour, A. (2013). Screening of snoring with an MP3 recorder. *Sleep & Breathing = Schlaf & Atmung*, 17(1), 77–84. doi:10.1007/s11325-012-0652-z
- LEE, J. B. O. K., PARK, Y. H., HONG, J. H. W. A., LEE, S. H., JUNG, K. I. H., KIM, J. E. H., ... SHIN, C. (2009). Determining optimal sleep position in patients with positional sleep-disordered breathing using response surface analysis. *Journal of Sleep Research*, 18(1), 26–35. doi:10.1111/j.1365-2869.2008.00703.x
- Lloberes, P., Levy, G., Descals, C., Sampol, G., Roca, A., Sagales, T., & de la Calzada, M. D. (2000). Self-reported sleepiness while driving as a risk factor for traffic accidents in patients with obstructive sleep apnoea syndrome and in non-apnoeic snorers. *Respiratory Medicine*, 94(10), 971–976. doi:10.1053/rmed.2000.0869
- Maimon, N., & Hanly, P. J. (2010). Does Snoring Intensity Correlate with the Severity of Obstructive Sleep Apnea? *Journal of Clinical Sleep Medicine*, 6(5), 475–478.
- Mallampati, S. R., Gatt, S. P., Gugino, L. D., Desai, S. P., Waraksa, B., Freiburger, D., & Liu, P. L. (1985). A clinical sign to predict difficult tracheal intubation; a prospective study. *Canadian Anaesthetists' Society*

- Masa, J. F., Corral, J., Pereira, R., Duran-Cantolla, J., Cabello, M., Hernández-Blasco, L., ... Montserrat, J. M. (2013). Effectiveness of sequential automatic-manual home respiratory polygraphy scoring. *The European Respiratory Journal*, 41(4), 879–87. doi:10.1183/09031936.00186811
- Matsiki, D., Deligianni, X., Vlachogianni-Daskalopoulou, E., & Hadjileontiadis, L. J. (2007). Wavelet-based analysis of nocturnal snoring in apneic patients undergoing polysomnography. *2007 Annual International Conference of the Ieee Engineering In Medicine and Biology Society, Vols 1-16*, IEEE Engn Med & Biol Soc. doi:10.1109/IEMBS.2007.4352690
- Mesquita, J., Sola-Soler, J., Fiz, J. A., Morera, J., & Jane, R. (2012). All night analysis of time interval between snores in subjects with sleep apnea hypopnea syndrome. *Medical & Biological Engineering & Computing*, 50(4), 373–381. doi:10.1007/s11517-012-0885-9
- Ministerio del Interior. Real Decreto 818 / 2009 , de 8 de mayo , por el que se aprueba el Reglamento General de Conductores, Pub. L. No. BOE-A-2009-9481 (2009). ESPAÑA: Boletín Oficial del Estado, Lunes 8 de junio de 2009. Retrieved from http://www.boe.es/diario_boe/verifica.php?c=BOE-A-2009-9481&acc=Verificar
- Montero Benavides, A., Blanco Murillo, J. L., Fernández Pozo, R., Espinoza Cuadros, F., Torre Toledano, D., Alcázar-Ramírez, J. D., & Hernández Gómez, L. A. (2015). Formant Frequencies and Bandwidths in Relation to Clinical Variables in an Obstructive Sleep Apnea Population. *Journal of Voice*, 1–9. doi:10.1016/j.jvoice.2015.01.006
- Montero Benavides, A., Fernández Pozo, R., Toledano, D. T., Blanco Murillo, J. L., López Gonzalo, E., & Hernández Gómez, L. (2014). Analysis of voice features related to obstructive sleep apnoea and their application in diagnosis support. *Computer Speech & Language*, 28(2), 434–452. doi:10.1016/j.csl.2013.08.002
- Oksenberg, A., & Silverberg, D. S. (1998). The effect of body posture on sleep-related breathing disorders: facts and therapeutic implications. *Sleep Medicine Reviews*, 2(3), 139–162. doi:10.1016/S1087-0792(98)90018-1
- Oppenheim, A., & Schaffer, R. (1968). Homomorphic analysis of speech. *IEEE Transactions on Audio and Electroacoustics*, 16(2), 221–226. doi:10.1109/TAU.1968.1161965
- Pack, Ai. I. (2012). *Sleep Apnea: Pathogenesis, Diagnosis and Treatment*. (Ai. I. Pack, Ed.) (2nd ed.). New York, NY 10017: Informa Healthcare USA, Inc. Retrieved from <http://books.google.com/books?hl=en&lr=&id=iMTB5MmyvAC&oi=fnd&pg=PP1&dq=Sleep+Apnea.+Pathogenesis,+Diagnosis+and+Tratment&ots=Se1ABt1mCd&sig=Lcsgni5hnGhcrxPZgQ7EIbUKh9U>
- Pevernagie, D., Aarts, R. M., & De Meyer, M. (2010). The acoustics of snoring. *Sleep Medicine Reviews*, 14(2), 131–144. doi:10.1016/j.smrv.2009.06.002
- Punjabi, N. M. (2008). The epidemiology of adult obstructive sleep apnea. *Proceedings of the American Thoracic Society*, 5(1546-3222 (Linking)), 136–143. doi:10.1513/pats.200709-155MG
- Reiter, M., Twerenbold, R., Reichlin, T., Benz, B., Haaf, P., Meissner, J., ... Mueller, C. (2012). Early diagnosis of acute myocardial infarction in patients with pre-existing coronary artery disease using more sensitive cardiac troponin assays. *European Heart Journal*, 33(8), 988–997. doi:10.1093/eurheartj/ehr376
- Roebuck, A., & Clifford, G. (2012). Multiscale Entropy Applied To Audio Data For Classifying Obstructive Sleep Apnoea Patients. In *B108. DIAGNOSTIC AND THERAPEUTIC APPROACHES IN SLEEP APNEA* (pp. A3841–A3841). American Thoracic Society. doi:10.1164/ajrcm-

- Roebuck, A., Monasterio, V., Geder, E., Osipov, M., Behar, J., Malhotra, A., ... Clifford, G. D. (2014). A review of signals used in sleep analysis. *Physiological Measurement*, 35(1), R1–57. doi:10.1088/0967-3334/35/1/R1
- Salmi, T. K., & Virtanen, V. (2009). FP37-WE-03 Innovation in sleep medicine: sleep disorder screening using mobile phone or MP3 recorder. *Journal of the Neurological Sciences*, 285, Suppl(0), S116 –. doi:10.1016/S0022-510X(09)70464-2
- Sanchez, A. I., Bermudez, M. P., & Buena-Casal, G. (2003). Short-term memory evaluation in sleep apnea patients before and after CPAP treatment. *Salud Mental*, 26(6), 55–61.
- Shete, D. S., & S.B. Patil, P. (2014). Zero crossing rate and Energy of the Speech Signal of Devanagari Script. *IOSR Journal of VLSI and Signal Processing*, 4(1), 01–05. doi:10.9790/4200-04110105
- Smith, J. O. (2002). Digital Audio Resampling Home Page. *Center for Computer Research in Music and Acoustics (CCRMA), Stanford University*. Retrieved May 29, 2015, from <http://www-ccrma.stanford.edu/~jos/resample/>
- Snell, R. C., & Milinazzo, F. (1993). Formant location from LPC analysis data. *IEEE Transactions on Speech and Audio Processing*, 1(2), 129–134. doi:10.1109/89.222882
- Snider, B. R., & Kain, A. (2013). Automatic Classification of Breathing Sounds During Sleep. *2013 IEEE International Conference On Acoustics, Speech and Signal Processing (icassp)*, Inst Elect & Elect Engineers; Inst Elect & Elect E.
- Sola-Soler, J., Jane, R., Fiz, J. A., & Morera, J. (2003). Spectral envelope analysis in snoring signals from simple snorers and patients with Obstructive Sleep Apnea. In *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No.03CH37439)* (Vol. 3, pp. 2527–2530). IEEE. doi:10.1109/IEMBS.2003.1280430
- Solé-Casals, J., Munteanu, C., Martín, O. C., Barbé, F., Queipo, C., Amilibia, J., & Durán-Cantolla, J. (2014). Detection of severe obstructive sleep apnea through voice analysis. *Applied Soft Computing*, 23, 346–354. doi:10.1016/j.asoc.2014.06.017
- van Maanen, J. P., Meester, K. A. W., Dun, L. N., Koutsourelakis, I., Witte, B. I., Laman, D. M., ... de Vries, N. (2013). The sleep position trainer: a new treatment for positional obstructive sleep apnoea. *Sleep & Breathing = Schlaf & Atmung*, 17(2), 771–9. doi:10.1007/s11325-012-0764-5
- Vlahandonis, A., Nixon, G. M., Davey, M. J., Walter, L. M., & Horne, R. S. C. (2013). Improvement of sleep-disordered breathing in children is associated with a reduction in overnight blood pressure. *Sleep Medicine*, 14(12), 1295–1303. doi:10.1016/j.sleep.2013.07.014
- Vlahandonis, A., Yiallourou, S. R., Sands, S. A., Nixon, G. M., Davey, M. J., Walter, L. M., & Horne, R. S. C. (2014). Long-term changes in heart rate variability in elementary school-aged children with sleep-disordered breathing. *Sleep Medicine*, 15(1), 76–82. doi:10.1016/j.sleep.2013.06.023
- Westerlund, A., Brandt, L., Harlid, R., Åkerstedt, T., & Trolle Lagerros, Y. (2014). Using the Karolinska Sleep Questionnaire to identify obstructive sleep apnea syndrome in a sleep clinic population. *Clinical Respiratory Journal*, 8(4), 444–454. doi:10.1111/crj.12095
- World Health Organization (WHO). (2015). Global Database on Body Mass Index - BMI classification. *Global Database on Body Mass Index*. Retrieved July 26, 2015, from http://apps.who.int/bmi/index.jsp?introPage=intro_3.html
- Yadollahi, A., & Moussavi, Z. (2010). Automatic breath and snore sounds classification from tracheal and ambient sounds recordings. *Medical Engineering & Physics*, 32(9), 985–990.

doi:10.1016/j.medengphy.2010.06.013

Glosario

AAC	Advanced Audio Coding
AAC-LC	AAC Low Complexity: Perfil de AAC de 16KBps a 448 KBps
AEPD	Agencia Española de Protección de Datos
API	Application Program Interface
AUC	Area Under the Curve
BBDD	Base de Datos
BMI	Body Mass Index
CBR	Constant Bit Rate
CSV	Fichero de valores separados por comas
ECG	Electrocardiograma
EEG	Electroencefalograma
EMG	Electromiograma
EOG	Electrooculograma
FFT	Fast Fourier Transform
FPR	False Positive rate
GB	Gigabyte
IMC	Índice de masa corporal (BMI en inglés)
Kbps	Kilobits por segundo
L.O.	Ley Orgánica
LTO	Linear Tape-Open. Tecnología de cinta magnética de almacenamiento de datos.
LTO-3	Generación 3 del formato LTO con 400 GB de almacenamiento nativo
LPC	Linear Prediction Coding
MAT	Fichero en formato propietario Matlab®
OMS	Organización Mundial de la Salud (WHO en inglés)
OSA	Obstructive Sleep Apnea
OSAHS	Obstructive Sleep Apnea-Hypopnea Syndrome
PCM	Pulse Code Modulation
PIB	Producto Interior Bruto
PSG	Polisomnografía
R.D.	Real Decreto
RAID	Redundant Array of Inexpensive Disks
RAID-5	RAID distribuido con paridad
RBF	Radial Basis Function
Res.	Respiración
ROC	Receiver Operating Characteristics
Ron.	Ronquido
SAHS	Síndrome de Apnea Hipo-apnea del Sueño
Sil.	Silencio
SMO	Sequential Minimal Optimization
SpO ₂	Nivel de saturación de oxígeno en sangre
STOP-BANG	Snore, Tired, Observed, blood Pressure – BMI, Age, Neck circumference, Gender
SVM	Support Vector Machine
TB	Terabyte
TFM	Trabajo Fin de Máster
TPR	True Positive Rate

USB	Universal Serial Bus
W.A.	Weighted Average
WAV	Apócope de Waveform Audio File Format
WEKA	Waikato Environment for Knowledge Analysis
WHO	World Health Organization (OMS en español)

Anexos

A. Base de datos de eventos respiratorios

Todos los datos recogidos y generados se han consolidado en una base de datos desarrollada sobre PostgreSQL 9¹⁴.

La decisión de utilizar esta base de datos se fundamenta en:

- Sistema Open Source y libre (en desarrollo desde 1995)
- Multiplataforma / Replicación
- Multiusuario
- Relacional y orientada a objetos
- Alta concurrencia
- Amplia variedad de tipos nativos. Manejo de Large Objects hasta 4TB.
- API para multitud de lenguajes
- Automatización del servidor en múltiples lenguajes

Se ha definido un esquema inicial (Figura A-1) donde se han registrado todos los datos obtenidos durante la prueba de concepto: condiciones del registro, datos de los voluntarios (anonimizados), datos procesados, atributos y clasificación.

Se ha verificado el acceso a los datos desde las herramientas WEKA¹⁵ y Matlab®¹⁶ mediante JDBC¹⁷ utilizando el conector de PostgreSQL¹⁸.

¹⁴ <http://postgresql.org>

¹⁵ <http://www.cs.waikato.ac.nz/ml/weka/>

¹⁶ <http://es.mathworks.com/products/matlab/>

¹⁷ <http://www.oracle.com/technetwork/java/javase/jdbc/index.html>

¹⁸ <http://jdbc.postgresql.org>

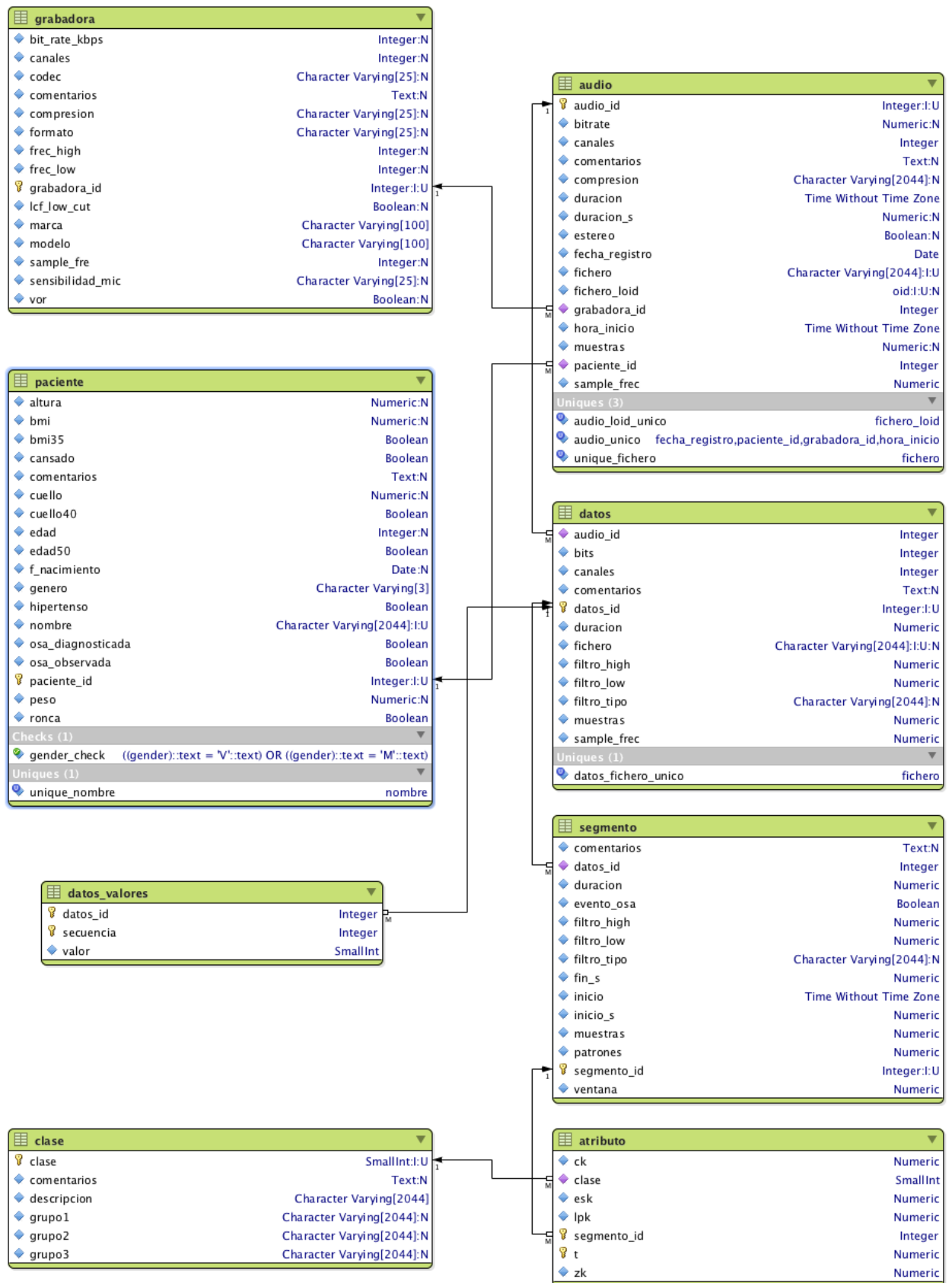


Figura A-1: Diagrama de la base de datos de ronquidos (v2.8)

Para facilitar la extracción de los datos, se ha creado una vista que realiza el JOIN de todas las tablas implicadas. La definición de la vista es la siguiente:

VISTA: ATRIBUTO_FULL

```
SELECT atributo.ck,
       atributo.esk,
       atributo.lpk,
       atributo.zk,
       (datos.sample_freq)::character varying AS sample_freq_datos,
       (audio.bitrate)::character varying AS bitrate,
       (audio.canales)::character varying AS canales_audio,
       audio.compresion AS comp,
       audio.sample_freq AS sample_freq_audio,
       (audio.grabadora_id)::character varying AS grabadora,
       grabadora.marca AS grabadora_marca,
       grabadora.modelo AS grabadora_modelo,
       paciente.altura,
       paciente.bmi,
       paciente.bmi35,
       paciente.cansado,
       paciente.cuello,
       paciente.cuello40,
       paciente.edad,
       paciente.edad50,
       paciente.genero,
       paciente.hipertenso,
       paciente.osa_diagnostics AS osa_diag,
       paciente.osa_observada AS osa_obs,
       paciente.peso,
       paciente.ronca,
       (paciente.paciente_id)::character varying AS paciente,
       clase.descripcion AS clase_l,
       clase.grupo1,
       clase.grupo2,
       clase.grupo3,
       (atributo.clase)::character varying AS clase_s,
       atributo.clase AS clase_n
FROM (((((atributo
       JOIN segmento ON ((atributo.segmento_id = segmento.segmento_id)))
       JOIN datos ON ((segmento.datos_id = datos.datos_id)))
       JOIN clase ON ((atributo.clase = clase.clase)))
       JOIN audio ON ((datos.audio_id = audio.audio_id)))
       JOIN paciente ON ((audio.paciente_id = paciente.paciente_id)))
       JOIN grabadora ON ((audio.grabadora_id = grabadora.grabadora_id)));
```


B. Funciones desarrolladas en Matlab® para acceso a la base de datos y cálculo de características

Función: extraer_segmento()

```
function extraer_segmento( segmento_id )
%EXTRAER_SEGMENTO Extrae de la BBDD de ronquidos los datos del segmento indicado
% Como salida genera un fichero WAV con el audio del segmento y un
% fichero MAT con los datos para su procesado. Crea todas las matrices
% necesarias para la obtención de características y las clases
% Calcula las características predefinidas
%
% Argumentos:
%   segmento_id : identificador del segmento a extraer (tabla segmento)
%
% Valores retornados:
%   ninguno

% Conectamos con la base de datos de ronquidos
conn = database('ronquidos','usuario','password',...
    'Vendor','PostgreSQL','Server','localhost');
set(conn,'AutoCommit','off');
setdbprefs('FetchInBatches','no')
setdbprefs('DataReturnFormat','numeric')

% Preparamos la cadena SELECT para obtener los metadatos del segmento
query = strcat('SELECT segmento.inicio_s, segmento.muestras, segmento.patrones,
segmento.ventana, datos.sample_frec, datos.datos_id FROM segmento INNER JOIN datos ON
segmento.datos_id = datos.datos_id WHERE segmento_id = ', num2str(segmento_id));
results = fetch(conn,query);

% Extraemos los metadatos del segmento en sus variables
inicio_s = results(1);
muestras = results(2);
patrones = results(3);
ventana = results(4);
sample_frec = results(5);
datos_id = results(6);
clear results
clear query

% Calculamos los límites del segmento dentro de la secuencia de datos
% Necesitamos 1 ms más para hacer la autocorrelación de la última ventana
inicio_m = inicio_s * sample_frec + 1;
final_m = round(inicio_m + muestras + (0.001 * sample_frec))-1;

% Preparamos la cadena SELECT para obtener los valores del segmento
query = strcat('SELECT valor FROM datos_valores WHERE (datos_id = ',num2str(datos_id),') AND
(secuencia BETWEEN (',num2str(inicio_m),') AND (',num2str(final_m),')) ORDER BY secuencia');
valores = fetch(conn,query);
clear query

% Cerramos la conexión con la base de datos
close(conn)
clear conn

% Normalizamos los valores de audio a [-1 +1], pues están en formato nativo
valores = double(valores-1)/32767.0;

% Guardamos el fichero de audio
fichero_wav = strcat('segmento_',num2str(segmento_id),'.wav');
audiowrite(fichero_wav,valores,sample_frec);

% Guardamos el fichero de datos
fichero_mat = strcat('segmento_',num2str(segmento_id),'.mat');
save(fichero_mat);
```

end

Función: extraer_segmentos_todos()

```
function extraer_segmentos_todos( datos_id )
%EXTRAER_SEGMENTOS_TODOS Extrae todos los segmentos de un registro de datos
% Extrae todos los segmentos de análisis definidos para un fichero de
% audio procesado y cargado en la tabla 'datos'
%
% Parámetros:
%   datos_id : identificador del fichero de audio en la tabla 'datos'
%
% Valores retornados:
%   ninguno

display(' ')
display(['Extrayendo todos los segmentos asociados al ID de datos ' num2str(datos_id)])
% Conectamos con la base de datos de ronquidos
conn = database('ronquidos','usuario','password',...
    'Vendor','PostgreSQL','Server','localhost');
set(conn,'AutoCommit','off');
setdbprefs('FetchInBatches','no')
setdbprefs('DataReturnFormat','numeric')

% Preparamos la cadena SELECT para obtener los identificadores de los
% segmentos asociados al registro de datos
query = strcat('SELECT segmento_id FROM segmento WHERE datos_id = ', num2str(datos_id));
segmentos = fetch(conn,query);

% Cerramos la conexión con la base de datos
close(conn)
clear conn

display(['Número de segmentos = ' num2str(size(segmentos,1))])
display(' ')
for s=1:size(segmentos,1)
    display(['Extrayendo segmento ' num2str(segmentos(s))])
    extraer_segmento(segmentos(s));
end
display(' ')
display('Extracción completada')
display(' ')

end
```

Función: analizar_segmento()

```
function analizar_segmento( segmento_id )
%ANALIZAR_SEGMENTO Analiza los datos del segmento indicado y calcula los
%valores de las características definidas para cada ventana
%
% Obtención de datos: cargar el fichero .mat normalizado con los datos
% de la señal de audio
%
% Sigo lo descrito en el artículo de Karunajeewa,2008.
%
% Características a extraer en cada ventana:
%   - Zk = Ratio de cruce por cero
%   - Esk = Energía de la señal
%   - Ck = Coeficiente de correlación a 1 ms de retardo
%   - Lpk = primer coeficiente predictor de LPC
%
% Una vez calculadas, guarda todos los resultados en el fichero .mat del
% segmento
```

```

%
% Argumentos:
% - segmento_id : número de segmento a procesar (tabla segmento)
%
% Valores retornados:
% ninguno
%
% Variables del fichero .mat del segmento:
% - inicio_s : segundo de inicio del segmento
% - muestras : número de muestras del segmento, incluido el retardo
% - patrones : número de ventanas a calcular (patrones)
% - ventana : ancho de la ventana de cálculo en segundos
% - sample_frec : frecuencia de muestreo
% - datos_id : identificador del fichero de audio del que se ha
% extraído el segmento
% - inicio_m : muestra en la que se inicia el segmento
% - final_m : muestra en la que finaliza el segmento
% - valores : datos de audio normalizados

% Cargamos el fichero .mat del segmento indicado (pasado como argumento)
fichero_mat = strcat('segmento_', num2str(segmento_id), '.mat');
display(' ')
display(['Leyendo fichero ' fichero_mat])
load(fichero_mat);

% Inicializamos los vectores de características y de clase
display('Inicializando vectores de características')
Zk = zeros(1,patrones);
Esk = zeros(1,patrones);
Ck = zeros(1,patrones);
Lpk = zeros(1,patrones);
clase = zeros(1,patrones); % Hay que añadirla a mano

% Calculamos las ventanas
mventana = floor(ventana * sample_frec); % muestras en la ventana
delay = floor(0.001 * sample_frec); % Muestras para el retardo

% Calculamos las características de cada ventana
display('Calculando características')
for k=0:patrones-1

    inicio = k*mventana+1;
    fin = k*mventana+mventana;

    x = valores(inicio:fin);

    % Calculamos el ZCR de la señal
    zcr = sum(abs(diff(x>=0)))/length(x);
    Zk(k+1) = zcr;

    % Calculamos la Energía media de la señal, en decibelios
    er = pow2db(sum(x.^2)/length(x));
    Esk(k+1) = er;

    % Coeficiente de correlación a lms de retardo
    xd = valores(inicio+delay:fin+delay);
    R = corrcoef(x,xd);
    Ck(k+1) = R(1,2);

    % Análisis LPC de 12 polos. Primer coeficiente predictor
    a = lpc(x,12);
    Lpk(k+1) = a(2);

end

% Tiempo de las ventanas/patrones
ts = (0:patrones-1)*(1/sample_frec)+inicio_s;

% Borramos las variables de cálculo intermedias
display('Housekeeping')
clear k inicio fin x zcr er xd R a mventana delay

% Guardamos el fichero .mat actualizado
display('Guardando datos actualizados del segmento')
save(fichero_mat);

display('Análisis terminado')
display(' ')

```

end

Función: subir_atributos()

```
function subir_atributos( id )
%UNTITLED Subir los atributos calculados de un segmento a la base de datos
% Cargar el fichero .mat con las características calculadas para el
% segmento especificado en la base de datos
%
% Argumentos:
% - id : identificador del segmento a cargar
%
% Valores devueltos:
% ninguno.
%
% El fichero .mat ya ha de tener calculados todos los campos a insertar en
% la base de datos de ronquidos.

% Cargamos el fichero .mat del segmento indicado (pasado como argumento)
% Cargamos solo las variables que vamos a insertar en la base de datos
fichero_mat = strcat('segmento_',num2str(id),'.mat');
display(' ')
display(['Leyendo fichero ' fichero_mat])
load(fichero_mat,'ts','Ck','Esk','Lpk','Zk','clase','patrones');

% Generamos la tabla de datos a insertar
display('Preparando tablas')
segmento_id = ones(patrones,1)*id;
data = [segmento_id ts' Ck' Lpk' Esk' Zk' clase'];
tablename = 'atributo';
colnames = {'segmento_id','t','ck','lpk','esk','zk','clase'};
data_table = array2table(data,'VariableNames',colnames);

% Conexión a la base de datos
display('Conexión a la base de datos')
conn = database('ronquidos','postgres','Calimero99',...
'Vendor','PostgreSQL','Server','localhost');

display('Insertando en la base de datos')
fastinsert(conn,tablename,colnames,data_table);

display('Cerrando conexión')
close(conn);

display('Proceso terminado')
display(' ')
```

end

Función: wav2csv()

```
function wav2csv( id,fichero_wav,fichero_csv )
%WAV2CSV Convertir WAV a CSV para subirlo a la BBDD
% Lee fichero WAV en formato nativo y genera un fichero
% CSV con el ID indicado y le número de secuencia de cada
% valor del fichero, para importarlo en la tabla datos_valores
%
% Argumentos:
% id : datos_id del registro asociado de la tabla datos
% fichero_wav : fichero WAV de entrada. Ha de ser MONO (un canal)
% fichero_csv : fichero de salida en formato CSV
```

```

display('Leyendo fichero')
[audio,Fs] = audioread(fichero_wav);
info = audioinfo(fichero_wav);

% Convertimos el audio normalizado a valores enteros
display('Convirtiendo audio')
valores = round((audio * 32767),0)+1;
clear audio

% Calculamos la secuencia de inserción
display('Generando secuencia')
registros = size(valores,1);
secuencia = (1:registros)';
display('Generando IDs')
datos_id = ones(registros,1)*id;

data = [secuencia datos_id valores];

clear datos_id secuencia valores;

display('Generando tabla')
tablename = 'datos_valores';
colnames = {'secuencia','datos_id','valor'};
data_table = array2table(data,'VariableNames',colnames);

display('Escribiendo fichero CSV con')
display(strcat(num2str(registros),' registros'))
writetable(data_table(1:end,:),fichero_csv,'Delimiter','');

display('Proceso terminado')

end

```

Función: extraer_atributos_segmento()

```

function extraer_atributos_segmento( segmento_id )
%EXTRAER_SEGMENTO_CLASE Extrae de la BBDD de ronquidos los datos del segmento indicado
% Como salida genera un fichero WAV con el audio del segmento y un
% fichero MAT con los datos del segmento ya procesado.
%
% Argumentos:
% segmento_id : identificador del segmento a extraer
%
% Valores retornados:
% ninguno

display(' ')
display(['extrayendo segmento ' num2str(segmento_id)])

% Conectamos con la base de datos de ronquidos
display('Conectando a la base de datos')
conn = database('ronquidos','postgres','Calimero99',...
'Vendor','PostgreSQL','Server','localhost');
set(conn,'AutoCommit','off');
setdbprefs('FetchInBatches','no')
setdbprefs('DataReturnFormat','numeric')

% Preparamos la cadena SELECT para obtener los metadatos del segmento
display('Recuperando datos de identificación')
query = strcat('SELECT segmento.inicio_s, segmento.muestras, segmento.patrones,
segmento.ventana, datos.sample_frec, datos.datos_id FROM segmento INNER JOIN datos ON
segmento.datos_id = datos.datos_id WHERE segmento_id = ', num2str(segmento_id));
results = fetch(conn,query);

% Extraemos los metadatos del segmento en sus variables
inicio_s = results(1);
muestras = results(2);
patrones = results(3);
ventana = results(4);
sample_frec = results(5);

```

```

datos_id = results(6);
clear results
clear query

% Obtenemos los atributos del segmento
display('Recuperando atributos del segmento')
query = strcat('SELECT ck FROM atributo WHERE (segmento_id = ',num2str(segmento_id),') ORDER
BY t ASC');
Ck = fetch(conn,query);
clear query
query = strcat('SELECT lpk FROM atributo WHERE (segmento_id = ',num2str(segmento_id),') ORDER
BY t ASC');
Lpk = fetch(conn,query);
clear query
query = strcat('SELECT esk FROM atributo WHERE (segmento_id = ',num2str(segmento_id),') ORDER
BY t ASC');
Esk = fetch(conn,query);
clear query
query = strcat('SELECT zk FROM atributo WHERE (segmento_id = ',num2str(segmento_id),') ORDER
BY t ASC');
Zk = fetch(conn,query);
clear query
query = strcat('SELECT clase FROM atributo WHERE (segmento_id = ',num2str(segmento_id),')
ORDER BY t ASC');
clase = fetch(conn,query);
% Calculamos los valores de tiempo para los atributos
%ta = ((0:patrones-1)*ventana+(ventana/2))+inicio_s;
ta = ((0:patrones-1)*ventana)+inicio_s;
clear query

% Calculamos los límites del segmento dentro de la secuencia de datos
% Necesitamos 1 ms más para hacer la autocorrelación de la última ventana
display('Recuperando datos de audio')
inicio_m = inicio_s * sample_freq + 1;
final_m = round(inicio_m + muestras + (0.001 * sample_freq))-1;

% Preparamos la cadena SELECT para obtener los valores del segmento
query = strcat('SELECT valor FROM datos_valores WHERE (datos_id = ',num2str(datos_id),') AND
(secuencia BETWEEN (' ,num2str(inicio_m),') AND (' ,num2str(final_m),')) ORDER BY secuencia');
valores = fetch(conn,query);
% Calculamos los valores de tiempo para el segmento
ts= (0:muestras-1)*(1/sample_freq)+inicio_s;
clear query

% Cerramos la conexión con la base de datos
close(conn)
clear conn

% Normalizamos los valores de audio a [-1 +1], pues estan en formato nativo
display('Normalizando audio')
valores = double(valores-1)/32767.0;

% Guardamos el fichero de audio
fichero_wav = strcat('segmento_',num2str(segmento_id),'.wav');
display(['Guardando fichero de audio ' fichero_wav])
audiowrite(fichero_wav,valores,sample_freq);

% Guardamos el fichero de datos
fichero_mat = strcat('segmento_',num2str(segmento_id),'.mat');
display(['Guardando fichero de datos ' fichero_mat])
save(fichero_mat);

display('Extracción terminada')
display(' ')
end

```


C. Ejemplo de formulario STOP-BANG



OHIO SLEEP MEDICINE INSTITUTE
CENTER OF SLEEP MEDICINE EXCELLENCE™

Main Office | 4975 Bradenton Avenue, Dublin Ohio 43017 | T 614.766.0773 | F 614.766.2599
Branch Office | 7277 Smith's Mill Rd., New Albany 43054 | T 614.775.6177 | F 614.775.6178

Name _____
Height _____ Weight _____
Age _____ Male / Female _____

STOP-BANG Sleep Apnea Questionnaire

Chung F et al Anesthesiology 2008 and BJA 2012

STOP		
Do you SNORE loudly (louder than talking or loud enough to be heard through closed doors)?	Yes	No
Do you often feel TIRED , fatigued, or sleepy during daytime?	Yes	No
Has anyone OBSERVED you stop breathing during your sleep?	Yes	No
Do you have or are you being treated for high blood PRESSURE ?	Yes	No

BANG		
BMI more than 35kg/m2?	Yes	No
AGE over 50 years old?	Yes	No
NECK circumference > 16 inches (40cm)?	Yes	No
GENDER : Male?	Yes	No

TOTAL SCORE		
-------------	--	--

High risk of OSA: Yes 5 - 8

Intermediate risk of OSA: Yes 3 - 4

Low risk of OSA: Yes 0 - 2



Figura C-1: Ejemplo de formulario STOP-BANG

D. Descripción de las clases de clasificadores de WEKA y sus parámetros utilizados en este trabajo

La información mostrada en este anexo se ha extraído de la ayuda y del manual de usuario de WEKA 3.6.12. Se presenta como ayuda para entender los parámetros utilizados en los distintos modelos de clasificadores, tal y como se reflejan en las tablas de resultados de las secciones 4.2 y 4.3.

CLASE: ZERO R

NAME

weka.classifiers.rules.ZeroR

SYNOPSIS

Class for building and using a 0-R classifier. Predicts the mean (for a numeric class) or the mode (for a nominal class).

OPTIONS

debug -- If set to true, classifier may output additional info to the console.

CLASE: NAIVE BAYES SIMPLE

NAME

weka.classifiers.bayes.NaiveBayesSimple

SYNOPSIS

Class for building and using a simple Naive Bayes classifier. Numeric attributes are modelled by a normal distribution.

For more information, see

Richard Duda, Peter Hart (1973). Pattern Classification and Scene Analysis. Wiley, New York.

OPTIONS

debug -- If set to true, classifier may output additional info to the console.

CLASE: LOGISTIC

NAME

weka.classifiers.functions.Logistic

SYNOPSIS

Class for building and using a multinomial logistic regression model with a ridge estimator.

There are some modifications, however, compared to the paper of leCessie and van Houwelingen(1992):

If there are k classes for n instances with m attributes, the parameter matrix B to be calculated will be an $m \times (k-1)$ matrix.

The probability for class j with the exception of the last class is

$$P_j(X_i) = \exp(X_i B_j) / ((\sum_{j=1..(k-1)} \exp(X_i B_j)) + 1)$$

The last class has probability

$$1 - (\sum_{j=1..(k-1)} P_j(X_i)) \\ = 1 / ((\sum_{j=1..(k-1)} \exp(X_i B_j)) + 1)$$

The (negative) multinomial log-likelihood is thus:

$$L = -\sum_{i=1..n} \{ \\ \sum_{j=1..(k-1)} (Y_{ij} * \ln(P_j(X_i))) \\ + (1 - (\sum_{j=1..(k-1)} Y_{ij})) \\ * \ln(1 - \sum_{j=1..(k-1)} P_j(X_i)) \\ \} + \text{ridge} * (B^2)$$

In order to find the matrix B for which L is minimised, a Quasi-Newton Method is used to search for the optimized values of the $m \times (k-1)$ variables. Note that before we use the optimization procedure, we 'squeeze' the matrix B into a $m \times (k-1)$ vector. For details of the optimization procedure, please check `weka.core.Optimization` class.

Although original Logistic Regression does not deal with instance weights, we modify the algorithm a little bit to handle the instance weights.

For more information see:

le Cessie, S., van Houwelingen, J.C. (1992). Ridge Estimators in Logistic Regression. *Applied Statistics*. 41(1):191-201.

Note: Missing values are replaced using a `ReplaceMissingValuesFilter`, and nominal attributes are transformed into numeric attributes using a `NominalToBinaryFilter`.

OPTIONS

`debug` -- Output debug information to the console.

`maxIts` (-M) -- Maximum number of iterations to perform.

`ridge` (-R) -- Set the Ridge value in the log-likelihood.

CLASE: MULTILAYERPERCEPTRON

NAME

`weka.classifiers.functions.MultilayerPerceptron`

SYNOPSIS

A Classifier that uses backpropagation to classify instances.

This network can be built by hand, created by an algorithm or both. The network can also be monitored and modified during training time. The nodes in this network are all sigmoid (except for when the class is numeric in which case the the output nodes become unthresholded linear units).

OPTIONS

GUI -- Brings up a gui interface. This will allow the pausing and altering of the neural network during training.

- * To add a node left click (this node will be automatically selected, ensure no other nodes were selected).*
- * To select a node left click on it either while no other node is selected or while holding down the control key (this toggles that node as being selected and not selected).*
- * To connect a node, first have the start node(s) selected, then click either the end node or on an empty space (this will create a new node that is connected with the selected nodes). The selection status of nodes will stay the same after the connection. (Note these are directed connections, also a connection between two nodes will not be established more than once and certain connections that are deemed to be invalid will not be made).*
- * To remove a connection select one of the connected node(s) in the connection and then right click the other node (it does not matter whether the node is the start or end the connection will be removed).*
- * To remove a node right click it while no other nodes (including it) are selected. (This will also remove all connections to it)*
- * To deselect a node either left click it while holding down control, or right click on empty space.*
- * The raw inputs are provided from the labels on the left.*
- * The red nodes are hidden layers.*
- * The orange nodes are the output nodes.*
- * The labels on the right show the class the output node represents. Note that with a numeric class the output node will automatically be made into an unthresholded linear unit.*

Alterations to the neural network can only be done while the network is not running, This also applies to the learning rate and other fields on the control panel.

- * You can accept the network as being finished at any time.*
- * The network is automatically paused at the beginning.*
- * There is a running indication of what epoch the network is up to and what the (rough) error for that epoch was (or for the validation if that is being used). Note that this error value is based on a network that changes as the value is computed. (also depending on whether the class is normalized will effect the error reported for numeric classes).*
- * Once the network is done it will pause again and either wait to be accepted or trained more.*

Note that if the gui is not set the network will not require any interaction.

autoBuild -- Adds and connects up hidden layers in the network.

debug -- If set to true, classifier may output additional info to the console.

decay -- This will cause the learning rate to decrease. This will divide the starting learning rate by the epoch number, to determine what the current learning rate should be. This may help to stop the network from diverging from the target output, as well as improve general performance. Note that the decaying learning rate will not be shown in the gui, only the original learning rate. If the learning rate is changed in the gui, this is treated as the starting learning rate.

hiddenLayers (-H) -- This defines the hidden layers of the neural network. This is a list of positive whole numbers. 1 for each hidden layer. Comma separated. To have no hidden layers put a single 0 here. This will only be used if autobuild is set. There are also wildcard values 'a' = (attribs + classes) / 2, 'i' = attribs, 'o' = classes, 't' = attribs + classes.

learningRate -- The amount the weights are updated.

momentum -- Momentum applied to the weights during updating.

nominalToBinaryFilter -- This will preprocess the instances with the filter. This could help improve performance if there are nominal attributes in the data.

normalizeAttributes -- This will normalize the attributes. This could help improve performance of the network. This is not reliant on the class being numeric. This will also normalize nominal attributes as well (after they have been run through the nominal to binary filter if that is in use) so that the nominal values are between -1 and 1

normalizeNumericClass -- This will normalize the class if it's numeric. This could help improve performance of the network, It normalizes the class to be between -1 and 1. Note that this is only internally, the output will be scaled back to the original range.

reset -- This will allow the network to reset with a lower learning rate. If the network diverges from the answer this will automatically reset the network with a lower learning rate and begin training again. This option is only available if the gui is not set. Note that if the network diverges but isn't allowed to reset it will fail the training process and return an error message.

seed -- Seed used to initialise the random number generator. Random numbers are used for setting the initial weights of the connections between nodes, and also for shuffling the training data.

trainingTime -- The number of epochs to train through. If the validation set is non-zero then it can terminate the network early

validationSetSize -- The percentage size of the validation set. (The training will continue until it is observed that the error on the validation set has been consistently getting worse, or if the training time is reached). If This is set to zero no validation set will be used and instead the network will train for the specified number of epochs.

validationThreshold -- Used to terminate validation testing. The value here dictates how many times in a row the validation set error can get worse before training is terminated.

CLASE: SMO

NAME

`weka.classifiers.functions.SMO`

SYNOPSIS

Implements John Platt's sequential minimal optimization algorithm for training a support vector classifier.

This implementation globally replaces all missing values and transforms nominal attributes into binary ones. It also normalizes all attributes by default. (In that case the coefficients in the output are based on the normalized data, not the original data --- this is important for interpreting the classifier.)

Multi-class problems are solved using pairwise classification (1-vs-1 and if logistic models are built pairwise coupling according to Hastie and Tibshirani, 1998).

To obtain proper probability estimates, use the option that fits logistic regression models to the outputs of the support vector machine. In the multi-class case the predicted probabilities are coupled using Hastie and Tibshirani's pairwise coupling method.

Note: for improved speed normalization should be turned off when operating on SparseInstances.

For more information on the SMO algorithm, see

J. Platt: Fast Training of Support Vector Machines using Sequential Minimal Optimization. In B. Schoelkopf and C. Burges and A. Smola, editors, Advances in Kernel Methods - Support Vector Learning, 1998.

S.S. Keerthi, S.K. Shevade, C. Bhattacharyya, K.R.K. Murthy (2001). Improvements to Platt's SMO Algorithm for SVM Classifier Design. Neural Computation. 13(3):637-649.

Trevor Hastie, Robert Tibshirani: Classification by Pairwise Coupling. In: Advances in Neural Information Processing Systems, 1998.

OPTIONS

buildLogisticModels -- Whether to fit logistic models to the outputs (for proper probability estimates).

c -- The complexity parameter C .

checksTurnedOff -- Turns time-consuming checks off - use with caution.

debug -- If set to true, classifier may output additional info to the console.

epsilon -- The epsilon for round-off error (shouldn't be changed).

filterType -- Determines how/if the data will be transformed.

kernel -- The kernel to use.

numFolds -- The number of folds for cross-validation used to generate training data for logistic models (-1 means use training data).

randomSeed -- Random number seed for the cross-validation.

toleranceParameter -- The tolerance parameter (shouldn't be changed).

CLASE: SMO.POLYKERNEL

NAME

weka.classifiers.functions.supportVector.PolyKernel

SYNOPSIS

The polynomial kernel : $K(x, y) = \langle x, y \rangle^p$ or $K(x, y) = (\langle x, y \rangle + 1)^p$

OPTIONS

cacheSize -- The size of the cache (a prime number), 0 for full cache and -1 to turn it off.

checksTurnedOff -- Turns time-consuming checks off - use with caution.

debug -- Turns on the output of debugging information.

exponent -- The exponent value.

useLowerOrder -- Whether to use lower-order terms.

CLASE: J48

NAME

weka.classifiers.trees.J48

SYNOPSIS

Class for generating a pruned or unpruned C4.5 decision tree. For more information, see

Ross Quinlan (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers, San Mateo, CA.

OPTIONS

binarySplits -- Whether to use binary splits on nominal attributes when building the trees.

confidenceFactor -- The confidence factor used for pruning (smaller values incur more pruning).

debug -- If set to true, classifier may output additional info to the console.

minNumObj -- The minimum number of instances per leaf.

numFolds -- Determines the amount of data used for reduced-error pruning. One fold is used for pruning, the rest for growing the tree.

reducedErrorPruning -- Whether reduced-error pruning is used instead of C.4.5 pruning.

saveInstanceData -- Whether to save the training data for visualization.

seed -- The seed used for randomizing the data when reduced-error pruning is used.

subtreeRaising -- Whether to consider the subtree raising operation when pruning.

unpruned -- Whether pruning is performed.

useLaplace -- Whether counts at leaves are smoothed based on Laplace.

CLASE: RANDOMFOREST

NAME

weka.classifiers.trees.RandomForest

SYNOPSIS

Class for constructing a forest of random trees.

For more information see:

Leo Breiman (2001). Random Forests. Machine Learning. 45(1):5-32.

OPTIONS

debug -- If set to true, classifier may output additional info to the console.

maxDepth -- The maximum depth of the trees, 0 for unlimited.

numFeatures -- The number of attributes to be used in random selection (see *RandomTree*).

numTrees -- The number of trees to be generated.

seed -- The random number seed to be used.

CLASE: RANDOMTREE

NAME

weka.classifiers.trees.RandomTree

SYNOPSIS

Class for constructing a tree that considers K randomly chosen attributes at each node. Performs no pruning. Also has an option to allow estimation of class probabilities based on a hold-out set (backfitting).

OPTIONS

KValue -- Sets the number of randomly chosen attributes. If 0, $\log_2(\text{number_of_attributes}) + 1$ is used.

allowUnclassifiedInstances -- Whether to allow unclassified instances.

debug -- If set to true, classifier may output additional info to the console.

maxDepth -- The maximum depth of the tree, 0 for unlimited.

minNum -- The minimum total weight of the instances in a leaf.

numFolds -- Determines the amount of data used for backfitting. One fold is used for backfitting, the rest for growing the tree. (Default: 0, no backfitting)

seed -- The random number seed used for selecting attributes.